

---

# Essays on Aggregation With Heterogeneous and Interacting Agents

---



## Inauguraldissertation

zur Erlangung des Doktorgrades der Fakultät Sozial- und Wirtschaftswissenschaften der  
Otto-Friedrich-Universität Bamberg.

*vorgelegt von*  
Jan Schulz-Gebhard.

Bamberg, 2023

# PROMOTIONSKOMMISSION

Professor Mishaël Milaković, PhD (Erstgutachter und Sprecher)

Professor Dr. Frank Westerhoff (Zweitgutachter)

Professor Dr. Christian Proaño (Drittgutachter)

Datum der mündlichen Prüfung: 20.03.2023

Dieses Werk ist als freie Onlineversion über das Forschungsinformationssystem (FIS; <https://fis.uni-bamberg.de>) der Universität Bamberg erreichbar. Das Werk steht unter der CC-Lizenz CC-BY mit Ausnahme des Kapitels 5 "A Network-Based Explanation of Inequality Perceptions". Für dieses gelten die Bestimmungen des deutschen Urheberrechts.

Lizenzvertrag: Creative Commons Namensnennung 4.0  
<http://creativecommons.org/licenses/by/4.0>



URN: urn:nbn:de:bvb:473-irb-593896

DOI: <https://doi.org/10.20378/irb-59389>



Diese Dissertation wurde zeitweise im Rahmen eines strukturierten Programms an der Otto-Friedrich-Universität Bamberg verfasst und materiell wie ideell durch die Hans-Böckler-Stiftung gefördert. Dafür bin ich sehr dankbar.

## Acknowledgements

First of all, I want to thank my supervisor Mishael Milaković for his continued support, guidance and encouragement in the last years. He and his distributional approach were the main reason for me to come to Bamberg for my Master's degree in 2015 and I will be forever immensely grateful for him showing me a novel way to make sense of the economic world that continues to fascinate and excite me. It is thanks to him that I was able to overcome the obstacles and moments of frustration pursuing a PhD tends to bring with it.

Furthermore, I want to express my heartfelt gratitude to Frank Westerhoff and Christian Proaño for always being open to provide feedback and questions and for their valuable and instructive comments in many Brownbag seminars and elsewhere as well as the Economics Department in Bamberg more broadly for providing the stimulating and welcoming environment that characterises Bamberg. Special thanks go to Alexander Hempfing for being a pillar of support and friendship in the last years as well as my colleagues at the chair, Philipp Mundt and Christian Babirat, for always having an open ear for me and the countless discussions ranging from nerdy historical anecdotes to the peculiarities of stochastic differential equations we had during the years. I am indebted to Lukas Eggersdorfer, Stephan Eitel, Max Gödl, Carsten Källner, Nicolas Klas, Marius Theiler and Eva Willer for their invaluable and able research assistance. I also want to thank my colleagues at the Behavioral Macroeconomics Doctoral Research Group for making my two years in the program a unique experience. It is due to the BaGBeM program and the discussions at the chair that I could channel my increasing frustration with the economics I learned during my undergraduate degree into something (hopefully) productive. In this context, I also want to express my sincere gratitude for the support - financial and non-financial - of the Hans-Böckler-Foundation in the first two years of my PhD, especially to Jens Becker and Iris Henkel. I also want to thank Björn Hacker who wrote several reports for the foundation on my behalf.

I am also grateful for the Bamberg Group for Pluralism in Economics as a place of spirited discussion in the best sense of the word and the source of many of my initial research ideas, especially Armin Aminian, Jan Denninger, Margitta Grötsch, Mark Kirstein, Christian Krumme, Kerstin Hötte, Leo Ipsen, Lasare Samartzidis, Max Stindt, Christina Summerer and Maybrit Wächter and all the other people I got to know through the German network and Bamberg group since we founded the group in Bamberg in 2015. The same holds for EAEPE, PKES, and INET YSI which continue to provide a much-needed forum for pluralist economics.

The most frequent co-author for the papers of this thesis is by far Daniel Mayerhoffer. I owe a lot to the inspiring and productive cooperation with him but am even more thankful

for his support and friendship for over a decade now. Who would have thought where our nightly discussions on pacifism in some pub in Bayreuth would lead to, 11 years later? (He is still wrong on that, by the way.) Similarly, I want to thank Jan Weber for the continued discussions and friendship across the Atlantic and the necessary international perspective on topics ranging from maximum entropy to the notion of competition, and always insightful remarks.

Last but certainly not least, I am immensely grateful for the unconditional love and support from my family, Anna and Maya Gebhard, and my parents, Susanna Stocker-Schulz and Wolfgang Schulz as well as from Miriam Bömer. Thanks for keeping me (mostly) sane in the last years and thanks for always making sure to remind me that there are other important things in the world apart from economics, distributions, and networks. And especially in the case of Anna, thank you for being as brilliant and extraordinary as you are, patiently continuing to read, comment on and discuss my unformed research ideas, and after all these years still dealing with the sloppy thinking of economists.

# Contents

<b>1</b>	<b>Preface and Synopsis</b>	<b>7</b>
	References . . . . .	12
<b>2</b>	<b>The Mesoeconomic Aggregation of Microeconomic Shocks</b>	<b>14</b>
2.1	Introduction . . . . .	16
2.2	Idiosyncratic Shocks and Mesoeconomic Aggregation . . . . .	19
2.3	Data and Estimation . . . . .	26
2.4	Results . . . . .	31
2.5	Discussion . . . . .	35
2.A	Appendix: Generic Residuals and Simple Linear Regression . . . . .	36
2.B	Appendix: Winsorizing, $\beta$ estimates and $R^2$ . . . . .	41
	References . . . . .	43
<b>3</b>	<b>How Wealthy Are the Rich?</b>	<b>47</b>
3.1	Introduction . . . . .	49
3.2	Model . . . . .	53
3.3	Data and Estimation . . . . .	55
3.4	Results . . . . .	58
3.5	Discussion . . . . .	67
3.A	Model . . . . .	69
3.B	Analytical Results for the Estimation of $\alpha$ under $q$ -Truncation . . . . .	71
3.C	Empirical Densities of Returns to Wealth Portfolios . . . . .	75
3.D	Total Wealth Estimates . . . . .	76
3.E	Equiprobable Sampling From a Power Law . . . . .	77
3.F	Logarithmic Sampling From a Power Law . . . . .	78
	References . . . . .	79
<b>4</b>	<b>Equal Chances, Unequal Outcomes? Network-Based Evolutionary Learning and the Industrial Dynamics of Superstar Firms</b>	<b>85</b>
4.1	Introduction . . . . .	87
4.2	Models of Evolutionary Learning as a Representation of Empirical Findings	89
4.3	Model Description . . . . .	94
4.4	Simulation Results . . . . .	98
4.5	Discussion . . . . .	108
	References . . . . .	112
<b>5</b>	<b>A Network-Based Explanation of Inequality Perceptions</b>	<b>120</b>
5.1	Introduction . . . . .	121
5.2	Related Literature . . . . .	123
5.3	Model . . . . .	128
5.4	Results . . . . .	135

---

5.5	Discussion . . . . .	147
5.A	Linkage Probabilities for Homophilic Networks . . . . .	149
5.B	Perceived Quantiles in Perception Networks . . . . .	151
5.C	Linkage Based on Logarithmised Income Differences . . . . .	154
	References . . . . .	157
<b>6</b>	<b>Perception and Privilege</b>	<b>165</b>
6.1	Introduction . . . . .	167
6.2	Related Literature . . . . .	168
6.3	Model . . . . .	170
6.4	Results . . . . .	177
6.5	Discussion . . . . .	186
6.A	Appendix: Calibrating the Composite Wage Gap . . . . .	189
	References . . . . .	191
<b>7</b>	<b>A Network Approach to Consumption</b>	<b>197</b>
7.1	Introduction . . . . .	199
7.2	Related Literature . . . . .	201
7.3	Model . . . . .	204
7.4	Results . . . . .	210
7.5	Discussion . . . . .	218
	References . . . . .	221
<b>8</b>	<b>Aggregate Consumption, Perception Networks and Functional Distribution</b>	<b>231</b>
8.1	Introduction . . . . .	233
8.2	Model . . . . .	234
8.3	Results . . . . .	241
8.4	Discussion . . . . .	250
	References . . . . .	253

# Preface and Synopsis

Aggregating the behaviour of individual agents at the micro level to explain macro level phenomena is at the heart of the social sciences. Aggregation becomes notoriously difficult, though, when agents are strongly heterogeneous and interact with each other, as shown for a wide range of fields such as consumer theory (Kirman 1992), production functions (Baqae and Farhi 2019; Felipe and McCombie 2005) or social choice theory (Arrow 1950, 1961). Rather than circumventing the aggregation problem by strong assumptions of additivity or representativeness, complexity economics embraces the heterogeneity and interaction of agents by using methods from interdisciplinary physics and numerical simulations (Arthur 2021). This is the preanalytic vision of this thesis. In particular, it focuses on the relevance of the meso level of aggregation for macroeconomic outcomes characterised by the interplay of distributional regularities and network interaction. Conceptually, the meso level is located between the usual micro- and macroeconomic levels of aggregation and describes inherently *relational* phenomena, i.e., relative sizes or interactions as formalised by networks. As I will argue throughout this thesis, accounting for a meso layer might drastically alter policy conclusions for wide set of phenomena ranging from consumer behaviour to aggregate fluctuations and inequality perceptions that might elude approaches with strong homogeneity assumptions.

The complexity perspective is in many ways complementary to the paradigm of behavioural macroeconomics that has proven so productive in recent years: While the focus of behavioural macroeconomics is to refine behavioural rules of agents at the micro level (cf. Gabaix 2020, for a recent excellent example), interaction and heterogeneity of those sophisticated agents can often only be considered to a very limited extent. At the other extreme, the behavioural assumptions of complexity approaches are often deliberately simple to enable a thorough investigation of heterogeneity and network interaction (El-sner 2010). Theoretical models with very parsimonious assumptions at the behavioural level but featuring complex interaction can explain phenomena that would otherwise require strong assumptions on the rationality of agents. An illustrative case is the concept of “rational irrationality” (Caplan 2000) often invoked to explain the empirically widespread biases in public perceptions on policy issues such as the share of the US federal budget put towards foreign aid. In this framework, misperceptions are explained by altering the utility function of agents to feature demand for ‘irrational beliefs’. By contrast, the meso

---

model in chapter 6 relies on correct perceptions within plausible social networks and does not require any kind of biased information processing or behavioural ingredients to replicate empirical misperceptions. Some attempts to put this preanalytic vision into practice constitute the main body of this thesis and provide a common theme for the topically disparate papers.

The second chapter embodies this preanalytic vision perhaps most strongly. In this chapter, my co-authors Anna Gebhard and Mishaël Milaković and I demonstrate that micro-level shocks matter even more for aggregate fluctuations than the literature on granularity (Gabaix 2011) and production networks (Acemoglu et al. 2012) taken separately suggests. We develop a parsimonious theoretical model to show that higher-order effects and network transmission are significant, whenever sales shares react to productivity shocks. Empirically, accounting for this meso-level interaction of heterogeneity in size and connectivity through supplier networks manifests itself in considerable performance gains when explaining macro variables. Our results can be used to identify systemically important industries, as it is only about 3% of them that statistically explain more than two thirds of the European business cycle. The empirical success of this first parsimonious modeling attempt also implies a potential for further synergies between the literature on distributional regularities that is largely focused on reduced form stochastic processes and the vast literature on production networks with full-fledged macro models. I hope to pick up on especially this potential synergy after this dissertation and examine the effects of interregional production networks and diversification (Bottazzi et al. 2019) afterwards.

The third chapter focuses on an entirely different meso-level structure – top tails of empirical wealth distributions – and examines, how different wealth accumulation processes at the micro level affect top wealth shares and tail inequality. As it turns out, the effect of scale-dependence in wealth accumulation is potentially massive and can lead to differences in estimated tail wealth reaching almost one order of magnitude. This is a disconcerting finding for empirical research on the super-rich since we also show that typical tests of scale-dependent accumulation might deliver false-positive artefacts caused by tiny non-response rates at the top that notoriously plague survey data. Since empirical research on wealth inequality has, in the absence of register data, to rely on such survey data, we also propose a sampling algorithm that exploits the functional form of power law-like regimes that might mitigate the problem. We illustrate our results by using a German rich list (by the *manager magazin*) and a popular German survey (the *Socio-Economic Panel*). Our findings indicate that wealth inequality and overall wealth levels in Germany are probably severely underestimated by conventional methods. The joint paper with my co-author Mishaël Milaković is already published as Schulz and Milaković (2023) in *Review of Income and Wealth*.

The fourth chapter “Equal chances, unequal outcomes?” also focuses on stochastically multiplicative accumulation but in the context of firm size distributions that empirically



---

follow Zipf’s law in the upper tail. We take a well-known agent-based model of industrial dynamics by Dosi et al. (2017) as our starting point that combines stochastically multiplicative learning with a replicator dynamics type of competition procedure. The baseline model replicates a plethora of stylised facts, including the Zipf distribution in the upper tail of empirical firm size distributions, exponential firm age distributions, and double-exponential size growth rate distributions. We generalise the model by introducing a network layer on which firms compete for local market power to model local, industry-specific productivity that includes the baseline model for a fully connected network as a special case. We show that the findings in the baseline model are not robust against this rather parsimonious extension and that the empirically well-established combination of Zipf’s law and superstar-like behaviour in firm growth only emerges for the knife-edge case of a fully connected network. Against ordoliberal orthodoxy, we also find that a level playing field in initial conditions does not guarantee an equitable distribution of sales or *realised* market power. In other words, equal starting opportunities do not lead to equal outcomes. Abstracting from the concrete model, we also show that Zipf’s law places rather strong restrictions on the stochastically multiplicative growth process: Initial conditions are not allowed to vary systematically with the specific *local* conditions of entrance. The joint paper with my co-author Daniel Mayerhoffer is already published as Schulz and Mayerhoffer (2021a) in the *Journal of Business Economics*.

The fifth chapter is concerned with *perceptions* of overall (income) inequality. The focus on perceptions results from the fact that beliefs and perceptions rather than necessarily the actual levels of inequality shape behaviour such as consumption and voting choices that we care about as social scientists. Empirically, perceptions are indeed strongly biased and do not at all coincide with actual levels of inequality. We show analytically and by means of numerical simulations that the known stylised facts in the pertinent literature emerge, when individuals form *unbiased* perceptions based on samples that are skewed by the network formation process. The only necessary ingredient is that tie formation in this network is homophilic in income, i.e., the empirically documented phenomenon that people tend to know people that have similar incomes to their own. Most notably, we show that the empirically documented ‘middle-class bias’ necessarily emerges within our model, replicating the stylised fact that almost everybody perceives themselves to occupy a middle position within the income distribution, irrespective of their actual income and income position. We thus show that apparently ‘irrational’ perceptions do not necessarily imply irrational information processing by agents. Our model also generates social networks that replicate salient features of empirical social networks, most notably, their small-world property. In this sense, the paper is arguably an example of how a parsimonious mesoeconomic model can successfully explain a battery of stylised facts and thus serve as an externally validated building block to investigate other domains, like the fifth, sixth, and seventh papers of this thesis show. The joint paper with my co-authors Daniel

---

Mayerhoffer and Anna Gebhard received the Egon-Matzner-Award for Socio-Economics 2022 by the Institute for Public Finance and Infrastructure Policy (IFIP) of TU Vienna and is already published as Schulz et al. (2022) in *Social Networks*.

The sixth chapter is concerned with perceptions of *intergroup* inequality along gendered and racialised lines. We build on the empirical finding by Shaikh et al. (2014) who show that wage distributions can be disaggregated according to gender or race with their separate distributions sustaining their exponential character. This allows us to characterise the full distribution just by the gender or racial wage gap and apply the homophilic network mechanism of Schulz et al. (2022) to the data. We show that perception formation in homophilic networks indeed replicates the empirically observed underestimation of empirical wage gaps but that this effect is quantitatively too strong. Calibration by a recent Israeli sample (Malul 2021) shows that women place a much higher weight on a correct global signal, i.e., the correct wage gap as e.g. communicated by the national media, and also have less biased perceptions of overall inequality. Being economically underprivileged thus paradoxically translates into epistemic privilege. Our findings also imply that educational campaigns on intergroup inequality are likely to succeed and that they should target the privileged. The joint paper with Daniel Mayerhoffer is already published as Mayerhoffer and Schulz (2022) in *Applied Network Science*.

The seventh chapter explores consumption decisions within social networks. Consumption is upward-looking, i.e., consumers try to “catch-up to the Joneses”, while the reference groups are determined by a perception network that is externally validated and replicates perceptions of inequality and empirical network topologies (cf. the fifth and sixth chapters). The main aim of the paper is to propose an alternative mechanism to the canonical approach based on intertemporal optimisation with respect to unobservable stochastic income streams and the Euler equation. We show that the meso-level interaction of agents in the perception network that are heterogeneous in their income within a parsimonious and purely static model is able to account for the observed stylised facts of empirical consumption expenditure distributions. Most notably, we can reconcile approximately log-normal expenditure distributions with distributions of current income that are decidedly not log-normal. This empirical regularity is typically taken as evidence for intertemporal optimisation (Battistin et al. 2009), while we show that the mismatch in distributional form can also emerge from static network interaction only. We also show that network segregation dampens aggregate consumption levels and mediates the effect of income inequality on consumption which hopefully sheds light on the micro-origins of the international differences in consumption expenditure patterns. The joint paper with Daniel Mayerhoffer is currently under review and was awarded the Herbert Simon Prize 2022 by the European Association for Evolutionary Political Economy (EAEPE). A preprint is available on the arXiv (Schulz and Mayerhoffer 2022), with an earlier version

---

being published as a working paper by the Bamberg Economic Research Group (BERG) as Schulz and Mayerhoffer (2021b).

The eighth chapter is single-authored and extends the sixth chapter by now looking at the *functional distribution* of income, i.e., the wage and the profit share of national income. I propose a parsimonious synthesis consumption model that introduces class-based differences in marginal propensities to consume (MPC) following Kalecki into a model of upward-looking status comparisons. The income distribution follows the empirically established two-class structure of an exponential bulk with a power law tail. This parsimonious extension of the traditional Kaleckian model has rather drastic aggregate consequences: I show that heterogeneous consumption propensities with lower MPCs out of profits are neither necessary nor sufficient to generate wage-led aggregate consumption. These findings might help to make sense of recent presumedly “anomalous” findings that consumption is empirically profit-led in some cases, contrasting the predictions of Kaleckian models. Network segregation tends to lower aggregate consumption and might even shift the regime of consumption from wage- to profit-led without requiring changes in the idiosyncratic MPCs of agents. Notably, I only recover the original Kaleckian intuition that redistribution in favour of wages boosts aggregate consumption if the consumption behaviour of capitalists and workers is sufficiently different from each other. These findings enable the possibility of *desegregation-led* growth with desegregationist measures both increasing aggregate consumption and shifting the consumption regime into a wage-led direction. My results also indicate that empirical studies on the wage-led/profit-led nexus should take *perceived inequality*, consumer norms as well as the population shares of both classes into account.

Each of the seven papers constituting chapters two to eight of this thesis are intended for independent publication (or, as indicated above, already published) and can thus be read without requiring prior knowledge from the other chapters. Yet, all of them are attempts at showing the relevance of the intermediate or meso-level step for aggregating from the micro to the macro level. Rather than providing an overarching framework of ‘meso’, my aim with this thesis is to complement those conceptual approaches (Holland 1987; Dopfer et al. 2004) and illustrate the usefulness of meso-scale approaches by providing specific examples. The findings of this thesis demonstrate that several canonical results need to be qualified when meso-level structures are accounted for (cf. e.g. purely granular theory or Kaleckian growth models).

Naturally, working on these phenomena has led to many follow-up questions: How do the empirical statistical regularities at the firm-level discussed in the fourth chapter relate to aggregate fluctuations and fit into the meso framework elaborated in chapter two? Can the methodology of chapter three developed for top tail wealth levels also be used to gauge the impact of extreme values for other distributions following Zipf’s law, e.g., the firm size distribution? How would agents vote on redistributive measures when confronted with

the skewed sampling procedures of chapters five and six? And, most importantly, given the advent of high quality social and economic network data (van der Laan et al. 2022): Can we empirically quantify the importance of the meso channel for aggregate outcomes? It is my sincere hope that I will be able to continue to work on these issues after finishing this thesis.

## References

- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016.
- Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346.
- Arrow, K. J. (1961). *Social choice and individual values*. Wiley: New York, 2nd edition.
- Arthur, W. B. (2021). Foundations of complexity economics. *Nature Reviews Physics*, 3(2):136–145.
- Baqaei, D. R. and Farhi, E. (2019). JEEA-FBBVA Lecture 2018: The Microeconomic Foundations of Aggregate Production Functions. *Journal of the European Economic Association*, 17(5):1337–1392.
- Battistin, E., Blundell, R., and Lewbel, A. (2009). Why is consumption more log normal than income? Gibrat’s law revisited. *Journal of Political Economy*, 117(6):1140–1154.
- Bottazzi, G., Li, L., and Secchi, A. (2019). Aggregate fluctuations and the distribution of firm growth rates. *Industrial and Corporate Change*, 28(3):635–656.
- Caplan, B. (2000). Rational irrationality: a framework for the neoclassical-behavioral debate. *Eastern Economic Journal*, 26(2):191–211.
- Dopfer, K., Foster, J., and Potts, J. (2004). Micro–meso–macro. *Journal of Evolutionary Economics*, 14(3):263–279.
- Dosi, G., Pereira, M. C., and Virgillito, M. E. (2017). The footprint of evolutionary processes of learning and selection upon the statistical properties of industrial dynamics. *Industrial and Corporate Change*, 26(2):187–210.
- Elsner, W. (2010). The process and a simple logic of ‘meso’. Emergence and the co-evolution of institutions and group size. *Journal of Evolutionary Economics*, 20(3):445–477.
- Felipe, J. and McCombie, J. S. (2005). How sound are the foundations of the aggregate production function? *Eastern Economic Journal*, 31(3):467–488.
- Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772.

- Gabaix, X. (2020). A Behavioral New Keynesian Model. *American Economic Review*, 110(8):2271–2327.
- Holland, S. (1987). *The market economy: From micro to mesoeconomics*. New York: St. Martin’s Press.
- Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives*, 6(2):117–136.
- Malul, M. (2021). (Mis)perceptions about the Gender Gap in the Labor Market. *Forum for Social Economics*, forthcoming.
- Mayerhoffer, D. M. and Schulz, J. (2022). Perception and privilege. *Applied Network Science*, 7(1):1–25.
- Schulz, J. and Mayerhoffer, D. M. (2021a). Equal chances, unequal outcomes? Network-based evolutionary learning and the industrial dynamics of superstar firms. *Journal of Business Economics*, 91(9):1357–1385.
- Schulz, J. and Mayerhoffer, D. M. (2021b). A network approach to consumption. *BERG Working Paper Series*, 173.
- Schulz, J. and Mayerhoffer, D. M. (2022). A network approach to consumption. *arXiv preprint*, 2203.14259.
- Schulz, J., Mayerhoffer, D. M., and Gebhard, A. (2022). A network-based explanation of inequality perceptions. *Social Networks*, 70:306–324.
- Schulz, J. and Milaković, M. (2023). How wealthy are the rich? *Review of Income and Wealth*, 69(1):100–123.
- Shaikh, A., Papanikolaou, N., and Wiener, N. (2014). Race, gender and the econophysics of income distribution in the USA. *Physica A: Statistical Mechanics and its Applications*, 415:54–60.
- van der Laan, J., de Jonge, E., Das, M., Te Riele, S., and Emery, T. (2022). A whole population network and its application for the social sciences. *European Sociological Review*, forthcoming.

# The Meso-economic Aggregation of Microeconomic Shocks

---

Joint work with A. Gebhard and M. Milaković.

We thank Simone Alfarano, Herbert Dawid, Sebastian Gechert, Alexander Hempfing, Tim Hagenhoff, Tomasz Makarewicz, Philipp Mundt, Christian Proaño and Frank Westerhoff as well as the participants of the Workshop on Heterogeneous Interacting Agents 2019 in London and the First Behavioral Macroeconomics Workshop in Bamberg for the helpful comments and the discussion. We also gratefully acknowledge the financial support by the Hans-Böckler-Foundation through grant PK045.

---

## Abstract

Recent theoretical advances have called the invariance of sales shares with respect to productivity differences into question, arguing that sizable higher-order macroeconomic effects may emerge when sales shares react to productivity shocks. Here we propose a parsimonious model for the adjustment of sales shares in response to productivity shocks, and operationalize it to readily quantify the impact of idiosyncratic shocks on aggregate fluctuations. Using input-output data for the European (EU28) economy, we find that our model significantly outperforms recently suggested specifications that rely either on the granularity of sales shares or the heavily skewed distribution of connections in production networks. While our results confirm earlier findings in the sense that microeconomic shocks are an important driver of macroeconomic fluctuations, we demonstrate that previous approaches substantially underestimate their relative impact because they fail to account for what we term the mesoeconomic interaction of granularity and network effects. Our main empirical finding is that idiosyncratic shocks to about three percent of industries already explain more than two thirds of the business cycle once we account for this interaction.

---

---

## Contents of this Paper

2.1	Introduction . . . . .	16
2.2	Idiosyncratic Shocks and Mesoeconomic Aggregation . . . . .	19
2.3	Data and Estimation . . . . .	26
2.3.1	Data . . . . .	26
2.3.2	Estimation . . . . .	28
2.4	Results . . . . .	31
2.4.1	Robustness: Macro Controls . . . . .	33
2.5	Discussion . . . . .	35
2.A	Appendix: Generic Residuals and Simple Linear Regression . . . . .	36
2.B	Appendix: Winsorizing, $\beta$ estimates and $R^2$ . . . . .	41
	References . . . . .	43

---

## 2.1 Introduction

We propose a parsimonious model for the perpetual adjustment of sales shares in response to productivity shocks between discrete periods that is readily estimated from available data. We demonstrate that previous empirical approaches substantially underestimate the aggregate impact of idiosyncratic shocks because they fail to account for (what we term) the mesoeconomic interaction of granularity and network effects. Using input-output data for the European (EU28) economy on the industry level, we find that shocks to less than three percent of industries already explain about two thirds of the business cycle when we account for mesoeconomic interaction.

Earlier theoretical advances have emphasized the importance of heavily skewed and fat-tailed distributions in firm size (Gabaix 2011) or in the connectivity of production networks (Acemoglu et al. 2012), arguing that idiosyncratic shocks on the microeconomic level can have considerable macroeconomic effects. These approaches essentially rest on the foundational theorem for efficient economies put forward by Hulten (1978), and mark an important deviation from the diversification argument of Lucas (1977), which claims that microeconomic shocks will cancel out in the aggregation process. In light of Hulten’s theorem, whereby the aggregate productivity of an economy can be expressed as a sales share-weighted sum of idiosyncratic productivity levels, empirical studies have typically assumed sales shares to be constant between periods of observation. Baqaee and Farhi (2019) call this invariance of sales shares with respect to productivity differences into question and show that sizeable higher-order effects can emerge when sales shares do respond to productivity shocks, and that the magnitude of such propagation effects will depend on the structure of the underlying production network. From this perspective, our contribution can also be viewed as an operationalization of Baqaee and Farhi’s theoretical considerations that explicitly accounts for the dynamical adjustment of sales shares. The rich framework Baqaee and Farhi employ allows for a very general application up to second-order macroeconomic impacts of productivity shocks. Empirical operationalizations are, however, hindered by the dimensionality of the problem that requires to specify micro-level production functions for all producers. The problem is compounded by the fact that the elasticities Baqaee and Farhi employ are equilibrium objects which cannot be empirically estimated in isolation. By contrast, our approach treats all micro and macro elasticities as well as input-linkages as homogeneous but allows for precise aggregation up to arbitrary order, fully characterized by observables. This does not only allow for straightforward empirical application, it also lets us interpret commonly employed estimation strategies in the literature economically within our framework. In a recent survey, Baqaee and Rubbo (2022) review the literature on aggregation within production networks and identify several channels going beyond Hulten’s theorem. They find two major reasons for relative sales share to vary with technology shocks and thus exact Domar



aggregation to fail: Inefficiencies in initial allocation and nonlinear shock responses. Our paper focuses on the latter and proposes a plausible micro propagation process enabling aggregation without the need to estimate price wedges and micro production functions. The aggregation from microeconomic shocks to macroeconomic behavior is conceptually not trivial because the interactions among economic agents can amplify or attenuate initial shocks. Consequently, borrowing a term from statistical physics, we adopt a *mesoscopic* perspective on aggregation (see, e.g., Imry 2008) and understand both the size distribution as well as the network structure as mesoeconomic objects that we use as intermediate layers for aggregation. Mesoeconomic objects are still subject to microeconomic forces, in our case the microeconomic elasticities of substitution, yet the properties emerging from the interaction of industries with each other are irreducible to their individual micro behavior. This irreducibility stems from the fact that relative sizes and network structures are not individual properties of the considered micro entities but can only be defined in relation to all other micro entities at once. While our model is obviously not the only conceivable approach to mesoeconomic aggregation, its empirical performance merits consideration and avoids many of the problems that direct aggregation from the micro to the macro level brings about.<sup>1</sup> Intuitively, shocks to firms or industries that are both large in size and central in the production network will have a substantial macroeconomic impact whereas shocks to small and peripheral entities will not. The mesoeconomic interaction of granularity and network effects is less intuitive when the two effects work in opposite directions. To gauge the relative importance of the two effects, we also derive the polar case of an economy that is homogeneous in its network structure yet heterogeneous in its size distribution (the purely granular case). The commonly used approximation by Riemann sums that leads to the purely granular case turns out to be a special case of our model when the network is regular. More importantly, though, the interaction of granularity and network structure substantially increases explanatory power compared to the purely granular special case.

Our paper relates to at least three strands of previous research. Hulten's theorem and Domar (1961) aggregation of idiosyncratic productivity shocks are the basis of modern growth accounting and its empirical applications, suggesting a linear impact of productivity shocks on aggregate fluctuations via sales shares (or Domar weights). The contributions by Gabaix (2011) and Acemoglu et al. (2012) are important because they demonstrate that even in a world of linear shock responses idiosyncratic shocks will have a significant impact on the business cycle if Domar weights are fat-tailed. Viewed from this perspective, their approaches are two sides of the same coin (Gabaix 2016), showing that idiosyncratic shocks will generally not cancel out in the process of aggregation. Second,

---

<sup>1</sup>Kirman (1992) argues that traditional aggregation rules have typically imposed an unrealistically large degree of homogeneity on economic agents, for instance the Gorman (1961) polar form, resulting in the inability of such models to explain the moments of aggregate output time series (Ascari et al. 2015).

from a microeconomic point of view, it is only the network perspective that sheds light on comovement and shock transmission among producers of intermediate goods, and microeconomic studies indeed find a large degree of sectoral comovement (Saint-Paul 1993; Foerster et al. 2011; Di Giovanni et al. 2014; Stella 2015), a property that we can also reproduce in our model. Third, Baqaee and Farhi (2019) challenge and qualify Hulten’s aggregation rule, extending it beyond first-order terms to allow for nonlinearities that can emerge when sales shares vary with productivity shocks. They argue that Hulten’s aggregation rule is an artifact of the constant returns to scale (CRS) property assumed for every micro production function, and recent sectoral level studies estimate elasticities of substitution that differ from the CRS benchmark (Atalay 2017; Boehm et al. 2019; Barrot and Sauvagnat 2016). This is also what casual empiricism suggests, since relative sales shares vary on all levels of aggregation from plants to firms to industries or sectors. To paraphrase Sims (1980), we face a “wilderness” of potential alternatives for the aggregation of production once we depart from the knife-edge Cobb-Douglas case, apparently preventing the estimation of the correct structural model. To circumvent this problem, we opt for a small set of nonparametric assumptions that imply restrictions on the functional form of the considered production function, resulting in a reduced-form model based on the elasticity of substitution among supplier relationships. Aggregate shocks are then measured by the mesoeconomically weighted sum of idiosyncratic productivity shocks, allowing us to decompose the aggregate shock into its constituent parts and to isolate the impact of each industry on aggregate fluctuations. Our mesoeconomic measure outperforms previous approximations that assume an invariance of sales shares between periods. It is worthwhile to point out that we achieve this performance using raw data only, that is without resorting to any arbitrary manipulation of the productivity growth rate distributions such as *winsorizing*, which considerably distorts the performance of granular approaches (Dosi et al. 2019).<sup>2</sup> Finally, our approach can also be thought of as a short-run complement to the complexity index proposed by Hidalgo and Hausmann (2009), who argue that the relative degree of complexity of national production structures is a useful predictor for national levels of production in the long run. Notably, though, our short-run perspective is transnational and emphasizes instead that aggregate short-run fluctuations depend on industry level characteristics that are not confined to national boundaries. The next section presents our reduced-form model. Section 3 deals with the econometric identification strategy for testing our model predictions. Our main empirical results are presented in section 4, including several robustness checks. We discuss the implications and limitations of our results in section 5.

---

<sup>2</sup>Since purely granular approaches to aggregate fluctuations are very sensitive to winsorizing, we quantify its impact for productivity shocks that have a normal or a double-exponential distribution in the Online Appendix C available at [https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi\\_lehrstuehle/vwl\\_internationale\\_wirtschaft/Online\\_Appendix\\_SGM.pdf](https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi_lehrstuehle/vwl_internationale_wirtschaft/Online_Appendix_SGM.pdf). We show analytically that in the empirically relevant leptokurtic case the bias from winsorizing is substantial.

## 2.2 Idiosyncratic Shocks and Mesoeconomic Aggregation

Hulten's theorem states that for any efficient economy in general equilibrium with  $N$  producers (firms or industries) indexed by  $i$

$$d \log TFP = \sum_{i=1}^N \frac{S_i}{Y} d \log TFP_i, \quad (2.1)$$

where  $TFP$  denotes the total factor productivity and  $Y$  the GDP of the economy, while  $S_i$  are the sales and  $TFP_i$  the productivity levels of entity  $i$ . Following Gabaix (2011) and assuming that GDP and TFP growth rates are related through a proportionality constant  $\mu$  reflecting factor usage, we have

$$d \log Y = \mu \sum_{i=1}^N \frac{S_i}{Y} d \log TFP_i. \quad (2.2)$$

This implies that for any given time period the rate of change of GDP is determined by the Domar weighted sum of idiosyncratic productivity shocks. Hulten's theorem holds for any efficient economy, irrespective of the existence or particular shape of input-output linkages in the economy, but merely up to a first-order approximation, or equivalently when the Domar weights  $S_i/Y$  are time invariant with respect to idiosyncratic productivity  $TFP_i$ . Whenever sales shares do respond to productivity shocks, however, potentially sizeable higher-order effects can emerge from these changes in sales shares. Baqaee and Farhi (2019) consider a second-order approximation and show that the impact of time-varying sales shares on GDP growth is uniquely determined by two effects: the derivatives of sales shares and of the input-output multiplier with respect to productivity shocks. While the former describes the change of sales shares with respect to productivity differences, the latter describes how these changes affect the ratio of total sales to GDP. It turns out that all higher order terms vanish only in the Cobb-Douglas CRS case. Whenever the production functions of some producers deviate from this knife-edge scenario, second-order terms can become quantitatively important, depending on the degree of factor reallocation that is assumed.

The operationalization of these theoretical insights is complicated by at least two concerns. First, the estimation of a complete structural model depends on unobservable general equilibrium elasticities of substitution. Second, since sales shares and productivity shocks are observable, we could construct a measure of the total change in TFP using (2.2) by integrating over a discrete time interval, where now both productivity levels  $TFP_i$  and

sales shares  $S_i/Y$  can vary over time,

$$\frac{\Delta Y_t}{Y_{t-1}} = \mu \sum_{i=1}^N \int_{t-1}^t \frac{S_i}{Y}(t) d \log TFP_i(t), \quad (2.3)$$

yet a good approximation of the integral in continuous time necessitates reasonably high observational frequencies. The discrete approximation by (left) Riemann sums that is used by both Gabaix (2011) and Baqaee and Farhi (2019) reads

$$\frac{\Delta Y_t}{Y_{t-1}} \approx \mu \sum_{i=1}^N \frac{S_{i,t-1}}{Y_{t-1}} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}, \quad (2.4)$$

and becomes more accurate the higher the frequency of observation. This approximation, however, presupposes that sales shares stay constant between  $t - 1$  and  $t$  and jump to their new level precisely at time  $t$ . Given quarterly or annual frequencies, it seems rather implausible for sales shares to stay (even approximately) constant between periods.<sup>3</sup>

Hence we want to improve and generalize approximation (2.4) in order to account for the adjustment in sales shares within a discrete time interval using observations only at  $t - 1$  and  $t$ . To do so, we employ the following parsimonious set of assumptions that recovers the approximation by Riemann sums in (2.4) as a special case. Since our data are on the industry level, the exposition will refer to industries, yet in principle it applies equally to firms. One assumption per level of aggregation fully determines our model, providing a convenient modular structure:

1. On the *micro* level, we follow standard procedure and assume that idiosyncratic productivity shocks are exogenous and that each industry's revenue and TFP growth rates are related by a finite proportionality constant  $\eta \geq 1$ . This condition is very general and is fulfilled by any conventional neoclassical production function on the micro level (Barro 1999).
2. On the *meso* level, we assume a constant and finite propagation intensity  $\omega > 1$  between supplying and supplied industries, implying that intermediate inputs are *micro substitutes* in the terminology of Baqaee and Farhi (2019).<sup>4</sup> If two industries do not have a supplier-recipient relationship, there is no direct effect from the growth of revenue in one to the other, but there will generally be an indirect effect that is mediated by the elasticity parameter.<sup>5</sup> Apart from this condition, we also

---

<sup>3</sup>While the input-output multiplier in our European data is remarkably stable over time, with an average value of around 0.5 and a coefficient of variation (CoV) of 0.01, the average CoV for Domar weights is more than one order magnitude larger at 0.2.

<sup>4</sup>We deliberately refrain from decomposing this substitutability in its partial price and quantity effect, since this would require us to define the degree of factor mobility, in particular labor mobility, which depends on the considered time-horizon (Baqaee and Farhi 2019).

<sup>5</sup>As a microfoundation for this incomplete pass-through, Liu (2020) consider (iceberg) adjustment costs. Coincidentally, with a similar but distinct model, they arrive at a very similar aggregation rule.

impose Gibrat's law in industry sizes and postulate that the size growth rate is uncorrelated to industry size. The rich extant literature on the firm level typically indicates that Gibrat's law is a reasonable approximation for the relationship of the mean growth rate and firm size but fails at higher moments (Sutton 1997, as the canonical reference). We confirm this finding for the industry level aggregation of our dataset, indicating that Gibrat's law for the first moment is indeed an innocuous assumption.<sup>6</sup>

3. On the *macro* level, we aggregate from end-of-period sales levels to GDP via a constant input-output multiplier  $\phi \in (0, 1]$  that measures the ratio of GDP to total sales, so  $\phi < 1$  implies the existence of intermediate goods, whereas  $\phi = 1$  corresponds to an islands economy without intermediate consumption.

Since we assume linear relations both at the micro and macro level, nonlinear effects have to originate at the meso level and it turns out that end-of-period sales levels are determined by the *interaction* of the size distribution with the topology of the production network, mediated by the propagation intensity.

Heuristically, the idea is to trace how a productivity shock to one industry propagates through the production network, and then to aggregate this effect over all industries experiencing simultaneous shocks. This will lead to an expression that resembles the centrality measure first proposed by Katz (1953), and it is also instrumental in deriving the purely granular and the purely networked economy as special cases. The linkages in the production network (or graph)  $G$  are described by the binary adjacency matrix  $A$ , where  $A_{ij} = 1$  if  $j$  is a supplier to  $i$  and  $A_{ij} = 0$  otherwise.<sup>7</sup> Let  $d$  denote the graph distance between two industries and consider what we term the *order effect* of a productivity shock to an industry  $j$  as a function of the graph distance to an industry  $i$  that is (possibly but not necessarily) different from  $j$ . Denote the rate of change in the sales of an industry  $j$  that has a distance  $d$  to the initially affected industry  $i$  by  $\Delta S_{j,t}^d / S_{j,t-1}$ . From our micro level assumption, the initial effect of a productivity shock on the sales of industry  $i$  is the zeroth-order effect denoted by  $\Delta S_{i,t}^0 / S_{i,t-1}$ , and is proportional to the initial productivity shock<sup>8</sup>

$$\frac{\Delta S_{i,t}^0}{S_{i,t-1}} = \eta \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \quad (2.5)$$

---

<sup>6</sup>Material available upon request.

<sup>7</sup>Note that we are deviating from convention here, with  $i$  typically being the supplier in standard IO tables. Our adjacency matrix is thus simply the transpose of the standard version.

<sup>8</sup>This is a conventional assumption in real business cycle models (see Gabaix 2011, for a more detailed discussion).

Suppose that each industry  $i$  takes goods produced by some industry  $j$  as inputs; given our meso level assumption of a constant propagation intensity  $\omega > 1$ , the first-order effect of a change in industry  $i$ 's sales on the sales growth of a supplier  $j$  to  $i$  is then given by

$$\frac{\Delta S_{i,t}^0}{S_{i,t-1}} = \omega \frac{\Delta S_{j,t}^1}{S_{j,t-1}}, \quad (2.6)$$

implying that the first-order effect on the sales growth of a supplier industry  $j$  stemming from a single productivity shock to industry  $i$  is<sup>9</sup>

$$\frac{\Delta S_{j,t}^1}{S_{j,t-1}} = \eta \omega^{-1} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \quad (2.7)$$

Next consider an industry  $k$  that is a supplier to industry  $j$ ; given a direct path from  $i$  to  $k$  through  $j$ , the second-order effect on  $k$  from the initial shock to  $i$  is therefore

$$\frac{\Delta S_{k,t}^2}{S_{k,t-1}} = \eta \omega^{-2} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}, \quad (2.8)$$

and so forth. It is convenient to rewrite equations (2.7) and (2.8) in terms of their order effect for an arbitrary industry  $j$  that might be but is not necessarily connected to the initially affected industry  $i$  by a distance  $d$ . This allows to generalize for an arbitrary number of paths between industries at a distance  $d$  by exploiting a property of the  $d$ -th power of the binary adjacency matrix  $A$ . For this, we note that  $A_{ij}^d$  denotes the number of paths between industry  $j$  and  $i$  through  $d - 1$  intermediaries. Thus  $A_{ij}^d$  gives the number of ways through which industry  $j$  is affected by a change in the productivity of industry  $i$  in the  $d$ -th round of adjustments, mediated by the constant propagation intensity.

If we assume that the network topology does not change during a discrete time interval (but might very well change from one period to the next), the  $n$ -th order effect on an arbitrary industry  $j$  is then given by

$$\frac{\Delta S_{j,t}^n}{S_{j,t-1}} = A_{ij}^n \eta \omega^{-n} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \quad (2.9)$$

Equation (2.9) and  $\omega > 1$  imply that the correlation between the sales growth rates of different industries has to be positive and decreasing in the network distance between industries, and this is exactly what Carvalho (2014) finds for his dataset of US industries, lending empirical support to the plausibility of this requirement. To get the total effect

---

<sup>9</sup>Note that  $\omega$  might vary between periods and could thus be indexed by  $t-1$ . We allow for time-varying weights in our empirical application, yet  $\omega$  is very stable in practice which is why we drop time indices not to clutter notation too much.

from the shock on  $i$  to an industry  $j$  from all orders, we sum over  $n$

$$\frac{\Delta S_{j,t}}{S_{j,t-1}} = \eta \sum_{n=0}^{\infty} A_{ij}^n \omega^{-n} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}, \quad (2.10)$$

and  $\omega > 1$  is a sufficient condition for this sum to converge. In the next step, we consider the total effect of simultaneous TFP shocks to all  $N$  industries apart from the singular one we considered so far and on their sales growth rates by summing both sides of the equation over the set of all industries by

$$\sum_{i=1}^N \frac{\Delta S_{i,t}}{S_{i,t-1}} = \eta \sum_{i=1}^N \sum_{j=1}^N \sum_{n=0}^{\infty} A_{ij}^n \omega^{-n} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \quad (2.11)$$

The expression in (2.11) closely resembles the Katz centrality vectors frequently employed in the network literature to measure the centrality of nodes when higher-order effects are relevant. To see this, consider the vector of Katz centrality scores  $k$  for the graph  $G$  with a binary but not necessarily symmetric adjacency matrix  $A$  depending on some attenuation parameter  $\beta > 1$  and given by

$$k(\beta) = ((\mathcal{I}_{N \times N} - \beta A)^{-1} - \mathcal{I}_{N \times N}) \mathcal{I}_{N \times 1} \quad (2.12)$$

$$= ((\mathcal{I}_{N \times N} + \beta^{-1} A + \beta^{-2} A^2 + \dots) - \mathcal{I}_{N \times N}) \mathcal{I}_{N \times 1} \quad (2.13)$$

$$= \left( \sum_{j=1}^N \sum_{n=0}^{\infty} \beta^{-n} A_{ij}^n - 1 \right)_{i=1}^N, \quad (2.14)$$

where  $\mathcal{I}_{N \times N}$  is the identity matrix with dimension  $N \times N$  and  $\mathcal{I}_{N \times 1}$  is a unit vector with length  $N$ . Bonacich (1991) shows that if  $\beta$  approaches the dominant (Perron) eigenvalue  $\lambda_{max}$  from above, the vector of Katz centrality scores approaches the vector of eigenvector centrality scores denoted  $e$ . For  $\beta = \omega$  we have from equation (2.14) that  $k(\omega) = (\sum_{j=1}^N \sum_{n=0}^{\infty} A_{ij}^n \omega^{-n} - 1)_{i=1}^N$ , so this term corresponds to the vector of Katz centrality scores. Assuming that  $\omega = \lambda_{max}$ , that is the coefficient  $\omega$  corresponds to the Perron eigenvalue, this implies that  $k(\lambda_{max}) = e$ , so the Katz centrality vector is equal to the vector of eigenvector centrality scores, and we have  $\mathcal{I}_{N \times 1} + e = (\sum_{j=1}^N \sum_{n=0}^{\infty} A_{ij}^n \omega^{-n})_{i=1}^N$ . Let  $\theta_{i,t-1} = 1 + e_{i,t-1}$  denote the  $i$ -th entry of the *shock transmission vector*  $\theta_{t-1}$ , capturing both the initial impact of the productivity shock as well as all meso level propagations between  $t$  and  $t - 1$ . Expressing equation (2.11) using  $\theta_{i,t-1}$ , it follows that

$$\sum_{i=1}^N \frac{\Delta S_{i,t}}{S_{i,t-1}} = \eta \sum_{i=1}^N \theta_{i,t-1} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \quad (2.15)$$

By imposing the restriction  $\omega = \lambda_{max}$ , the total rate of change resulting from the sum of arbitrary order-effects between  $t$  and  $t - 1$  will correspond to the centrality weighted sum of all productivity shocks. Hence we have characterized the end-of-period state of

all industries in the economy, which completes the meso level aggregation. To close the system and to aggregate to the macro level, we recall from our macro assumption that

$$\phi \sum_{i=1}^N \Delta S_{i,t} = \Delta Y_t, \quad (2.16)$$

where  $\Delta Y_t$  denotes the change in GDP between  $t-1$  and  $t$ , and the input-output multiplier  $\phi < 1$  indicates the presence of intermediate goods. From equation (2.16) we have

$$\Delta Y_t = \phi \sum_{i=1}^N \Delta S_{i,t} = \phi \sum_{i=1}^N \frac{\Delta S_{i,t}}{S_{i,t-1}} S_{i,t-1}, \quad (2.17)$$

and therefore

$$\frac{\Delta Y_t}{Y_{t-1}} = \phi \sum_{i=1}^N \frac{\Delta S_{i,t}}{S_{i,t-1}} \frac{S_{i,t-1}}{Y_{t-1}}. \quad (2.18)$$

To combine the meso condition in equation (2.15) with the macro level condition in equation (2.18), we need to impose the statistical regularity of Gibrat's law, i.e., industry size measures  $(S_{i,t-1}/Y_{t-1})_{i=1}^N$  and size growth rates  $(\Delta S_{i,t}/S_{i,t-1})_{i=1}^N$  being uncorrelated. It proves instructive to consider both vectors as sequences of realizations of random variables with finite first moment to rewrite equation (2.18) in terms of expectation values by

$$\phi \sum_{i=1}^N \frac{\Delta S_{i,t}}{S_{i,t-1}} \frac{S_{i,t-1}}{Y_{t-1}} = \phi N \cdot E\left[\frac{\Delta S_{i,t}}{S_{i,t-1}} \cdot \frac{S_{i,t-1}}{Y_{t-1}}\right] = \phi N \cdot E\left[\frac{\Delta S_{i,t}}{S_{i,t-1}}\right] \cdot E\left[\frac{S_{i,t-1}}{Y_{t-1}}\right], \quad (2.19)$$

with the last step exploiting Gibrat's law. This leads to

$$\phi N \cdot E\left[\frac{\Delta S_{i,t}}{S_{i,t-1}}\right] \cdot E\left[\frac{S_{i,t-1}}{Y_{t-1}}\right] = \phi \sum_{i=1}^N \frac{\Delta S_{i,t}}{S_{i,t-1}} \cdot E\left[\frac{S_{i,t-1}}{Y_{t-1}}\right], \quad (2.20)$$

and by equation (2.15)

$$\phi \eta \sum_{i=1}^N \theta_{i,t-1} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}} \cdot E\left[\frac{S_{i,t-1}}{Y_{t-1}}\right] = \phi \eta \sum_{i=1}^N \theta_{i,t-1} \frac{S_{i,t-1}}{Y_{t-1}} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}, \quad (2.21)$$

where the last equality assumes independence of Domar weights and weighted TFP shocks.<sup>10</sup> This allows us finally to restate the GDP growth rate as a function of industry level weights and idiosyncratic TFP shocks by

$$\frac{\Delta Y_t}{Y_{t-1}} = \phi \eta \sum_{i=1}^N \theta_{i,t-1} \frac{S_{i,t-1}}{Y_{t-1}} \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \quad (2.22)$$

---

<sup>10</sup>Note that this latter independence assumption is approximately true for our dataset, with weighted TFP shocks and Domar weights being only weakly connected on average with  $\rho \approx 0.1$ .



Equation (2.22) is our central result and relates the macro state of the economy to micro productivity shocks through the mesoeconomic interaction of network structure and Domar weights. We obtain equation (2.22) under the crucial assumption of micro substitutability ( $\omega > 1$ ) and the technical requirement that the parameter  $\omega$  is equal to the Perron eigenvalue  $\lambda_{max}$ . While the former is theoretically intuitive if we want to exclude explosive growth, the latter is certainly plausible from an empirical point of view as the Perron eigenvalue is remarkably stable over time and close to twenty, implying that about five percent of productivity shocks are transmitted to suppliers. We also invoke the statistical requirement of Gibrat's law that we can empirically confirm for our dataset. Note that the Riemann sum approximation (2.4) is a special case of our model specification. We can interpret this case economically as the correct specification for homogeneous network topologies or, more precisely,  $k$ -regular graphs (not merely as a discrete approximation with desirable mathematical properties). For  $k$ -regular graphs, all industries exhibit equal (or representative) degree  $k \geq 0$ . For this condition,  $\theta_{i,t-1} = 1 + \bar{e} = \bar{\theta} \geq 1$  for all  $i = 1, \dots, N$ . Letting  $\mu = \phi \eta \bar{\theta}$ , this leaves us with the formulation used by Gabaix (2011) and Baqaee and Farhi (2019),

$$\frac{\Delta Y_t}{Y_{t-1}} = \phi \eta \bar{\theta} \sum_{i=1}^N \frac{S_{i,t-1}}{Y_{t-1}} \cdot \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}} = \mu \sum_{i=1}^N \frac{S_{i,t-1}}{Y_{t-1}} \cdot \frac{\Delta TFP_{i,t}}{TFP_{i,t-1}}. \quad (2.23)$$

Even though the result in (2.23) is formally equivalent to (2.4) used in Gabaix (2011),  $\mu$  depends in our formulation not only on factor usage but also on the input multiplier as well as the average connectivity of producers. This result might help to explain why empirical studies after Gabaix' seminal contribution tend to find values for  $\mu$  far below unity (Ebeke and Eklou 2017; Miranda-Pinto and Shen 2019; Wagner and Weche 2020), as do we, in obvious violation to the interpretation by Gabaix (2011) in terms of different growth models which unanimously imply  $\mu > 1$ . The invariance of sales shares here is essentially a network homogeneity result. This alternative derivation is in line with the theoretical result in Baqaee and Farhi (2019) that the invariance of sales shares is a linearity result, since asymmetries in the network topology are the only possible source of nonlinearities here. We include this purely granular specification as a robustness check in the empirical part. Given the modular outline of the model, the purely granular case corresponds to a direct aggregation from the micro to the macro level without any heterogeneity at the meso level. In addition to a readily testable aggregation rule, the modular structure of our reduced-form model thus provides a convenient way to quantify the relevance of the meso-level step in the aggregation.

## 2.3 Data and Estimation

We use the second release of the World input-output database (2017), henceforth WIOD, that presently covers the period from 2000 to 2014. A detailed breakdown of the WIOD and its various data sources, for the most part based on national accounts, is provided by Timmer et al. (2015). Macroeconomic data for GDP and GDP per capita series as well as several other macro variables used for robustness checks are taken from the World Bank’s Development Indicators database (World Bank 2018). We focus on the EU28, mainly to exploit the large degree of cross-country connectivity in an integrated free trade area with a large ensemble of industries. While the sample period is rather short, it contains at least two important crises periods (the global financial and economic crisis and the European sovereign debt crisis) with substantial variation in GDP growth rates. Thus the sample period provides a reasonably challenging testing ground for the empirical modeling of aggregate fluctuations in general, and it should also help in discriminating between two different cases we have derived in the previous section.

### 2.3.1 Data

The WIOD provides annual series for 56 industries per country, leaving us with 1568 distinct industries in the EU28. The size-based metrics are taken from the Socio-Economic Accounts (SEA), in particular the annual volume of industry sales (gross output) and the number of employees used to calculate Domar weights and productivity shocks, respectively. The complementary cumulative distribution functions (CDFs) of industry sales, reported for each sample year in the Online Appendix D available at [https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi\\_lehrstuehle/vwl\\_internationale\\_wirtschaft/Online\\_Appendix\\_SGM.pdf](https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi_lehrstuehle/vwl_internationale_wirtschaft/Online_Appendix_SGM.pdf), reveal a large degree of heterogeneity in industry size, but they do not exhibit power law-like behavior. Thus aggregation from the firm to the industry level apparently prevents the straightforward application of ‘pure’ granular theory that rests on the power law property of firm size distributions.<sup>11</sup>

In addition to size-based metrics from the SEA, our model also requires the specification of adjacency matrices in order to account for the dynamic adjustment of sales shares through supplier relations. To construct these matrices, we binarize the World Input-Output Tables (WIOTs) from the WIOD that describe inter-industry flows for all industry

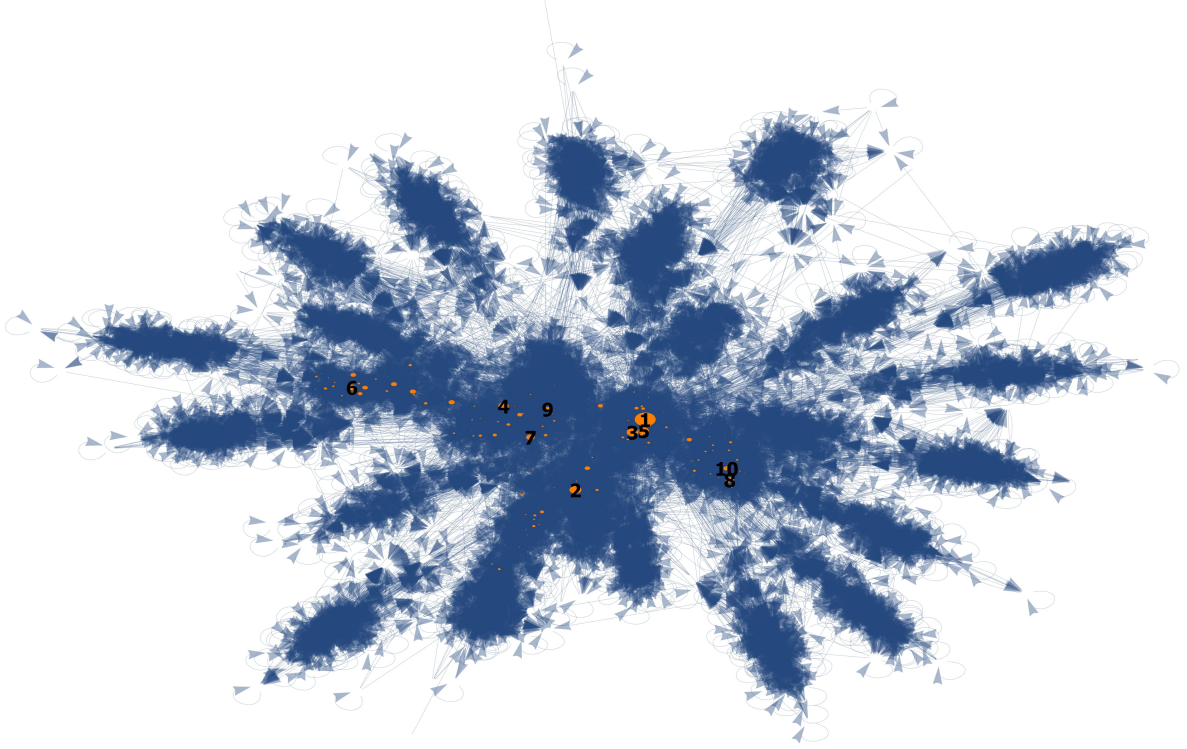
---

<sup>11</sup>By contrast, the distribution of our construct of ‘meso weights’ summarized in Online Appendix D available at [https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi\\_lehrstuehle/vwl\\_internationale\\_wirtschaft/Online\\_Appendix\\_SGM.pdf](https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi_lehrstuehle/vwl_internationale_wirtschaft/Online_Appendix_SGM.pdf) exhibits much larger variability in all years and is closer to the power law benchmark. In particular, the meso weights span approximately two orders of magnitude for the whole sample period which Stumpf and Porter (2012) considers to be a minimum requirement of power law behavior, while the granular weights fail to meet this minimum condition in any year.

pairs, leading to a  $1568 \times 1568$  matrix for each year. Following Carvalho (2014), we apply a standard threshold of one percent of an industry's total input deliveries for binarization. The resulting networks are extremely sparse and visualized in Figure 2.1 for the year 2000, where we observe 26 924 non-zero edges out of a possible  $1568^2 \approx 2.5 \times 10^6$ , resulting in a low network density of around 0.01, with an average degree of 17 in that year. Since almost all industries are connected through a (weak) input linkage, this leads to a massive reduction in average connectivity by almost two orders of magnitude, from 90 % of non-zero edges in the initial network to a mere 1% after binarization. This discrepancy leaves us reasonably confident that we effectively separated noise from signal and that the remaining links represent economically meaningful linkages. The size of each node in the figure is proportional to its meso weight score,<sup>12</sup> and the network looks very similar in other years as well. After all, the binarization threshold does not artificially inflate connectivity and should consequently not have a spurious impact on aggregate fluctuations. Yet mesoeconomic aggregation emphasizes the importance of higher-order effects in the production network, and the figure shows that all clusters are still connected to each other via some intermediaries, corroborating the importance of shock transmission through higher-order connections in spite of the network's low density and average degree. The industry with highest impact on aggregate fluctuations is the German automotive sector, in line with the intuition that industries producing highly complex final products rely on long and diversified supply chains and are thus also of systemic importance.

---

<sup>12</sup>In the empirical part the transmission vector corresponds to eigenvector centralities, which are much more efficiently computed by a standard power iteration method than Katz centralities. This does not bias the estimates in our dataset because the initial impact of productivity shocks rapidly vanishes in relation to the higher-order terms in equation (2.11), even for a finite-sum approximation.



- |  |   |
|--|---|
| 1 German manufacture of motor vehicles, trailers and semi-trailers | 6 Swedish real estate   |
| 2 French construction  | 7 British human health and social work activities                       |
| 3 German construction  | 8 Italian construction  |
| 4 British construction   | 9 British Public administration and defence; compulsory social security |
| 5 German manufacture of machinery and equipment n.e.c.             | 10 Italian wholesale trade  |

Figure 2.1: The EU28 production network exhibits substantial heterogeneity in meso weights and connectivity, along with clusters that mostly coincide with national borders. The ten industries with highest weights are listed in descending order and node sizes are proportional to meso weights.

### 2.3.2 Estimation

To quantify the impact of industry level shocks on aggregate fluctuations we essentially adhere to the procedure of Gabaix (2011). First we construct *residuals* that represent the sum of idiosyncratic shocks to the  $K < N$  largest industries, here weighted according to the theoretically motivated specifications in equations (2.22) and (2.23).<sup>13</sup> Then we run linear regressions on the various residuals and consider the coefficients of determination, in other words their respective (adjusted)  $R^2$ , in order to compare how much of the variation in GDP growth can be explained by each of the two model specifications. To do so, we need to measure productivity shocks and here we will use labor productivity instead of TFP.<sup>14</sup> Let  $g_{i,t} - \bar{g}_t$  denote the productivity shock to industry  $i$  at time  $t$ , where

<sup>13</sup>The granular case studied by Gabaix corresponds to eq. (2.23), but here for industries instead of firms.

<sup>14</sup>Our model is built on TFP shocks that are Hicks-neutral, while labor productivity shocks generally are not. This is at least one of the reasons why the growth of TFP and labor productivity follow

$g_{i,t} = \ln(S_{i,t}/E_{i,t}) - \ln(S_{i,t-1}/E_{i,t-1})$  is the observed change in  $i$ 's labor productivity, with  $E_{i,t}$  being the number of employees in industry  $i$  at time  $t$ , while  $\bar{g}_t$  measures the aggregate productivity shock. This conventional formulation assumes that the observed productivity signal  $g_{i,t}$  is additively separable into an idiosyncratic and an aggregate component. We calculate the aggregate shock following Gabaix (2011) by taking the median over the  $K^*$  industries with highest weight in each period. The specification of  $\bar{g}_t$  thus leaves one additional degree of freedom - the number of industries  $K^*$  with highest weight over which we calculate the aggregate shock. Our results remain remarkably robust for several different  $K^*$  that we use for our calibration.

Now we are in a position to compute residuals for our respective model specifications, starting with the *mesoeconomic residual*  $\mathcal{M}$  that aggregates the idiosyncratic productivity shocks to the  $K$  largest industries  $i = 1, \dots, K$  in period  $t$  using the mesoeconomic weight  $\theta_{i,t-1} \cdot S_{i,t-1}/Y_{t-1}$ ,

$$\mathcal{M}_t = \sum_{i=1}^K \theta_{i,t-1} \frac{S_{i,t-1}}{Y_{t-1}} (g_{i,t} - \bar{g}_t). \quad (2.24)$$

For robustness we also consider the case of purely linear shock responses on a  $k$ -regular graph economy with the *granular residual*  $\mathcal{G}$ , using merely Domar weights

$$\mathcal{G}_t = \sum_{i=1}^K \frac{S_{i,t-1}}{Y_{t-1}} (g_{i,t} - \bar{g}_t), \quad (2.25)$$

The mesoeconomic model and its granular variant translate to the following regressions,

$$I : (a) \ g_t^Y = \beta_0 + \beta_1 \mathcal{M}_t + \epsilon_t \quad (b) \ g_t^Y = \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 \mathcal{M}_{t-1} + \epsilon_t \quad (2.26)$$

$$II : (a) \ g_t^Y = \beta_0 + \beta_1 \mathcal{G}_t + \epsilon_t \quad (b) \ g_t^Y = \beta_0 + \beta_1 \mathcal{G}_t + \beta_2 \mathcal{G}_{t-1} + \epsilon_t \quad (2.27)$$

where aggregate fluctuations are denoted  $g_t^Y$  and measured by the growth rates of (per capita real) GDP. The left panel, marked (a), corresponds to equations (2.22) and (2.23) while the right panel, marked (b), includes a lagged term of the respective residual. The lagged term accounts for the possibility that the adjustment of sales shares takes time such that GDP responds to micro productivity shocks with a delay, and controls for reverse causality.

---

different trajectories, as shown for instance by Syverson (2004) for the manufacturing industry on the plant level. Nevertheless we employ labor productivity for three reasons: from a theoretical viewpoint, the estimation of TFP presupposes precisely the CRS assumption we challenge with our model; from an empirical viewpoint, we simply do not have TFP data in the WIOD; finally, the use of labor productivities facilitates the comparison with the granular results of Gabaix, who also theorizes with TFP but then utilizes labor productivity for his empirical exercise.

A useful tool for our investigation are plots that show how much of the explained variation  $R^2$  in GDP growth rates can be attributed to the largest  $K$  industries.<sup>15</sup> To establish a benchmark we show in Appendix 2.A that the expected  $R^2$  in the presence of idiosyncratic shocks, denoted  $E[R^2]$ , for  $K$  industries with weights  $w \in \{\theta S/Y, S/Y\}$  is given by the ratio of Hirschman-Herfindahl indices  $H$  for the weights of these  $K$  industries relative to the weights for all  $N$  industries,

$$E[R^2] = \frac{H(K)}{H(N)} = \sum_{i=1}^K \frac{(w_i)^2}{\sum_{j=1}^N (w_j)^2}. \quad (2.28)$$

Since both weights exhibit considerable leptokurtosis, we should observe a large initial increase in  $R^2$  already for a small number of industries, followed by a roughly monotonous yet moderate increase that settles on a plateau when  $K$  approaches its upper bound, where the size of industry weights decreases rapidly. These three qualitative features will guide our comparison and assessment of the different specifications in the next section.

Notably, and in stark contrast to conventional granular approaches (Gabaix 2011; Ebeke and Eklou 2017; Blanco-Arroyo et al. 2018), we only use raw data in all our specifications, and do not winsorize the empirical density of productivity growth rates.<sup>16</sup> In Appendix 2.B, we show that winsorizing should actually *decrease* rather than increase explanatory power for a correct model specification, essentially due to a decrease in residual variance, and this effect is amplified by the well-documented leptokurtic nature of empirical productivity growth rate densities (Dosi et al. 2012; Yu et al. 2015; Dosi et al. 2019).<sup>17</sup> Explanatory power thus only increases when the winsorizing procedure generates spurious comovement between industries, inflating regression coefficients by compensating for the initial drop in variance and explanatory power. Viewed from this perspective, the fact that a purely granular approach needs to generate comovement through winsorizing to push the estimates over the brink of significance already indicates the crucial importance of mesoeconomic interactions.

<sup>15</sup>This approach is inspired by Blanco-Arroyo et al. (2018), who investigate purely granular regimes, though. Here we generalize the comparison of empirical  $R^2$  plots to the benchmark of correctly identified idiosyncratic shocks and weights.

<sup>16</sup>Winsorizing is typically justified by reference to extraordinary events such as large mergers. Yet Bottazzi et al. (2019) show in their appendix that for each of the idiosyncratic events causing the four minimum and maximum growth rates in the Gabaix (2011) sample, the notion of “extraordinary events” can be reasonably challenged.

<sup>17</sup>We show that symmetric winsorizing at merely one percent already diminishes the explanatory power by more than five percent when productivity growth rates follow a Laplace distribution, as the empirical literature strongly suggests. The loss in explanatory power is considerably smaller for a Gaussian distribution at about 1.8 percent. We discuss the details of these computations in Appendix 2.B and Online Appendix C available at [https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi\\_lehrstuehle/vwl\\_internationale\\_wirtschaft/Online\\_Appendix\\_SGM.pdf](https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi_lehrstuehle/vwl_internationale_wirtschaft/Online_Appendix_SGM.pdf). Recent studies also recognize the relevance of leptokurtic productivity growth rate distributions (Carvalho and Tahbaz-Salehi 2019; Daniele and Stüber 2020).

## 2.4 Results

The empirical formulation of our main idea corresponds to eq. (2.26), so we estimate eq. I(a), along with the variant I(b) that includes one lagged term of the mesoeconomic residual. We set the number of relevant industries  $K^*$  to  $K^* = 200$  but our results remain robust for different  $K^*$  and  $K$ .<sup>18</sup> Figure 2.2 shows our main comparative results for the mesoeconomic and granular residual for all different specifications. The mesoeconomic specifications lead to highly significant contemporaneous coefficient estimates for both considered specifications, while coefficients for the granular special case remain insignificant. The mesoeconomic specification accounting for higher order effects leads to considerable performance gains to the purely granular case, lending robust support to the mesoeconomic hypothesis. The explanatory power of the mesoeconomic specification as measured by the coefficient of determination is sizeable, more than one half or more than two thirds of explained variance for the purely contemporaneous and specification with lagged term, respectively. Our findings do not appear to be driven by misspecification or anomalous behavior either, with all coefficients exhibiting the expected (positive) sign<sup>19</sup> and estimated intercepts displaying very similar magnitude (of  $\approx 10^{-2}$ ) and thus not artificially inflating model performance or biasing results in favor of one residual. The time series of both residuals against GDP growth rates suggest that the superior explanatory power of the mesoeconomic residual is primarily the result of its performance during the twin crises within our sample period that the meso residual tracks reasonably well but are not at all reflected in the behavior of the purely granular variant. This invites to speculate that the higher-order effects the meso specification is designed to capture are particularly relevant to account for cascading failures that are presumedly most impactful in crisis periods.

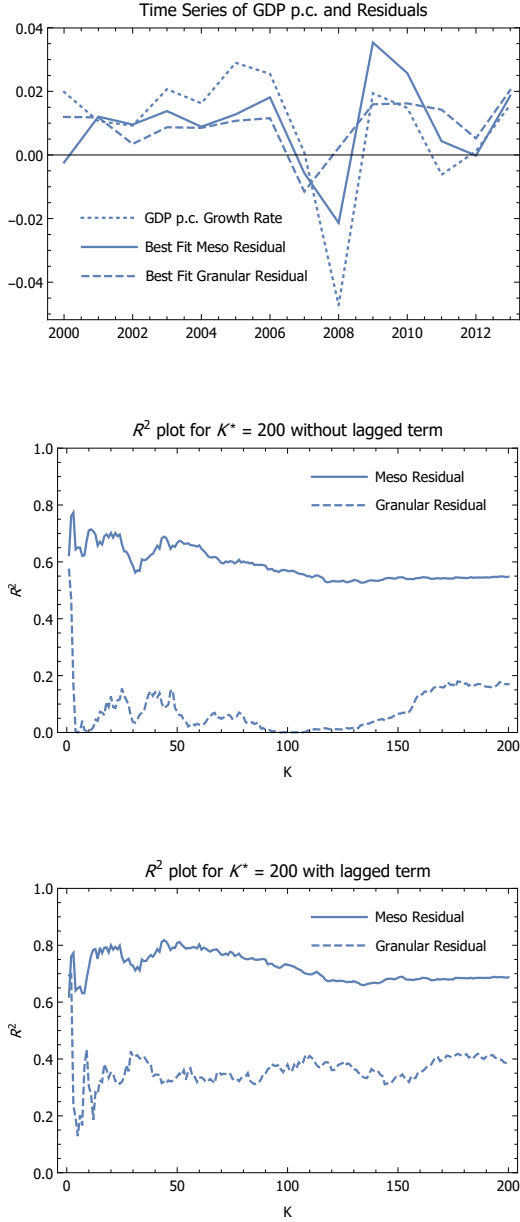
The incremental gains of explanatory power by adding industries are extremely asymmetric for the meso case.<sup>20</sup> The  $R^2$  plot quickly stabilizes after only about 50 industries - or 3% of all with systemic importance and thus follows our theoretical predictions on the functional form of  $R^2$  plots with leptokurtic weight distributions. The  $R^2$  plots furthermore indicate that the ‘snapshot’ results we report in the right panel of Figure 2.2 are not misleading and the coefficients of determination are robust against incremental variation of  $K$ . Most importantly, the meso residual outperforms the granular special case for the whole range of  $K = 1, \dots, K^*$ , demonstrating the importance of mesoeconomic shock

---

<sup>18</sup>Material available upon request.

<sup>19</sup>Since the eigenvector centrality we employ can, by the nature of eigenvectors, give centrality scores only to up to a common factor, both coefficient estimates are not directly quantitatively interpretable. We can thus gauge the empirical success of the mesoeconomic hypothesis only by the sign and significance of the estimated coefficients.

<sup>20</sup>Note that we take the composite  $K$  largest industries, e.g. when  $K = 10$  we take the ten largest industries in a given year so their identities might change over time.



I		
	(a)	(b)
$\mathcal{M}_t$	5.984*** (1.569)	6.761*** (1.456)
$\mathcal{M}_{t-1}$		1.376 (1.441)
Intercept	0.010** (0.004)	0.009** (0.003)
T	14	13
$R^2$	0.55	0.69
Adjusted $R^2$	0.51	0.62
F Statistic	14.54***	11.00***
II		
	(a)	(b)
$\mathcal{G}_t$	0.697 (0.445)	0.608 (0.418)
$\mathcal{G}_{t-1}$		0.871* (0.456)
Intercept	0.009* (0.005)	0.010* (0.005)
T	14	13
$R^2$	0.17	0.39
Adjusted $R^2$	0.10	0.26
F Statistic	2.45	3.148*

Figure 2.2: The mesoeconomic residual outperforms the granular special case for all specifications and all considered  $K^*$ . The left panel shows the time series of the meso and granular residuals against  $g_t^Y$  as well as the respective  $R^2$  plots without and with lagged term. The right panel reports the corresponding regression results for all specifications. Standard errors are given in parentheses; significance at the \*\*\* 1%, \*\* 5%, \* 10% level.



transmission for aggregate fluctuations.<sup>21</sup> The lagged terms change the overall picture only marginally. The granular variant gains relatively more from including a lagged term but is still robustly inferior to the mesoeconomic variant. Both contemporaneous coefficients are also rather stable across specifications and remain highly significant. The lagged term in the meso specification adds little to the explanatory power of the meso residual and is insignificant at the usual significance levels. In accordance with the mesoeconomic hypothesis, it is thus primarily contemporaneous developments in the mesoeconomic residual that appears to comove with aggregate fluctuations. By contrast, Domar-aggregated shocks in the granular special case fail to explain much of the variation in GDP growth rates in both variants. Estimates from conventional Domar aggregation thus appear to strongly underestimate the impact of microeconomic idiosyncratic shocks. Since purely granular economies are a special case of the general mesoeconomic aggregation for fully symmetric networks, this finding also provides indirect evidence for the relevance of accounting for *asymmetric* topologies, with only a few industries occupying central roles within the production network.

### 2.4.1 Robustness: Macro Controls

Our investigation of potential confounding variables starts with two exogenous aggregate disturbances, i.e., the rate of change in oil prices (as oil is close to a universal input) and monetary policy decisions, in order to exclude the possibility that the performance of the mesoeconomic model stems from aggregate shocks that might unwittingly be included in our (rather conscientious) construction of idiosyncratic shocks. To preserve a reasonable number of degrees of freedom, we consider only the contemporaneous versions of the control variables in our specifications. We use the following control variables, following Gabaix (2011):

- (i) ECB policy rate (denoted ECB, in percent, Euro area) and
- (ii) growth rate of the real crude oil price (denoted Oil, Brent Europe).

These give rise to the following specifications:

$$\begin{aligned}
 (A) : g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 \text{ECB}_t + \epsilon_t \\
 (B) : g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 \text{Oil}_t + \epsilon_t \\
 (C) : g_t^Y &= \beta_0 + \beta_1 \mathcal{M}_t + \beta_2 \text{ECB}_t + \beta_3 \text{Oil}_t + \epsilon_t
 \end{aligned}$$

Each specification (A) to (C) is estimated by OLS. The results are summarized in Table 2.1. The  $\hat{\beta}_1$  coefficient stays highly significant across all specifications and, more

---

<sup>21</sup>This result is robust for variation in  $K^* = 50, 100, 150, 250$ . We provide these results upon request.

	(Benchmark)	(A)	(B)	(C)
$\mathcal{M}_t$	5.984*** (1.569)	5.999*** (1.243)	4.738** (1.949)	5.445*** (1.63)
ECB <sub>control</sub>		0.006** (0.002)		0.006** (0.002)
Oil <sub>control</sub>			0.020 (0.019)	0.009 (0.017)
Intercept	0.010** (0.004)	-0.006 (0.006)	0.007 (0.004)	-0.006 (0.006)
$R^2$	0.55	0.73	0.59	0.75
Adjusted $R^2$	0.51	0.69	0.52	0.67
AIC	-77.41	-83.15	-76.74	-81.58
BIC	-75.50	-80.60	-74.25	-78.38
F-Statistic	14.54***	15.64***	7.92***	9.87***

Table 2.1: The table summarizes the various robustness checks to examine the variability and performance of our benchmark mesoeconomic model against the inclusion of two macro controls (indicated by the subscript “control”). The low variability of effect sizes as well as the fact that the meso coefficient does not lose significance for any control indicates the robustness of the benchmark specification. The performance only improves when adding a simple proxy for monetary policy that our baseline model cannot account for per construction.

importantly, effect sizes do not vary much for the inclusion of all of the different controls (A) to (C). The largest deviation occurs for the specification that includes the rate of change in oil prices, but even this drop of about a quarter in effect size is rather negligible in size. Our previous estimated effect size therefore appears to be robust against the inclusion of standard control variables, indicating that the model performance is not driven by omitted confounders.

The specifications in which we include a simple proxy for monetary policy ((A) and (C)) would be preferred to our benchmark model according to both information criteria.<sup>22</sup> This result, however, is hardly surprising since the mesoeconomic residual is derived in a Real Business Cycle (RBC) setting where monetary policy is ineffective in generating real effects (Snowdon et al. 1994). As e.g. Galí (2015) and the whole New Keynesian literature demonstrate, interest rate shocks that are transmitted through the amount of available credit, inflation or asset prices have potentially strong effects on GDP fluctuations under price-stickiness. In this sense, complementing the real-side analysis with simple proxies

<sup>22</sup>As Kuha (2004) notes, both criteria select for different types of data-generating processes. The BIC is more suitable for processes with a small number of strong effects, while the AIC selects better for models with a collection of many effects. Model selection without such prior information should therefore be based on both criteria in agreement.

for monetary policy shocks can be expected to outperform a model built only on real-side effects.<sup>23</sup>

By contrast, sudden surges in oil prices were one of the major historical motivations for supply-side shocks within RBC-type models, since oil price developments can (for economies like the EU28) be reasonably characterized as exogenous disturbances affecting production (Summers 1986). Our results leave little room for exogenous aggregate shocks in the form of oil price changes, though, when the microeconomic shocks to individual industries are properly accounted for within the meso residual. These results therefore not only testify to the general robustness of the mesoeconomic channel, they also suggest that conventional aggregate energy price shocks are efficiently identified by microeconomic shocks to a vanishingly small number of industries with disproportionate aggregate influence. Our findings thus corroborate the baseline assumption of RBC models in the sense that exogenous productivity shocks are the most important determinant of aggregate fluctuations in the short run but also show that they need to be conceived as operating on a much more granular level.

## 2.5 Discussion

Since Hulten (1978), growth theorists explicitly or implicitly believed to live in a world of linear shock responses. Our parsimonious model and the empirical results for Europe, however, cast doubt on the conventional wisdom that economies are linear in this sense, and thus on traditional Domar aggregation. We show that it is the strongly non-linear *interaction* of both size- and network-based channels, in the extant literature so far being treated entirely separately, that provides a crucial missing link between idiosyncratic productivity developments and aggregate GDP movement. This interaction channel emerges naturally, whenever we consider the dynamics between observation periods. Our reduced form model explicitly includes the adjustment of sales shares but still allows immediate Domar-type empirical application.

In contrast to much of contemporary macroeconomic theory, we do not counterfactually impose homogeneity in size or independence between industries to enable aggregation. Instead, we explicitly acknowledge industries as strongly heterogeneous and interdependent entities. Conveniently, this also allows us to make full use of the available information in the WIOD, which manifests itself in considerable performance gains. As we expect in a model without nominal variables, we find that the performance increases even further by including a simple proxy for monetary policy. Perhaps unexpectedly, however, we find

---

<sup>23</sup>This proxy can, by construction, not account for the (increasingly relevant) unconventional policy measures by the ECB and does only apply to the subset of the 19 Euro zone countries for our EU28 sample. The superior performance from including this imperfect proxy thus points to the relevance of the monetary policy channel and highlights its negligence in RBC-type models even more.

that the relative explanatory power of traditional aggregate shocks in the form of changes in oil prices is rather low compared to the mesoeconomic residual.

The robustly superior performance of the mesoeconomic aggregation rule against the rival granular special case suggests that idiosyncratic shocks are of much greater relevance than previously found and can account for about two thirds of aggregate fluctuations. Notably, we find that the aggregate relevance of shocks is extremely skewed towards those pertaining to merely 50 systemically relevant industries. In the European context of our paper, the proverbial notion of an ‘asymmetric shock’ therefore lives on in another form, now not with shocks to countries but to specific industries (De Grauwe 2020). In analogy to established results in complex network theory, we find the system of European industries to be resilient against equiprobable adverse shocks, as the systemically relevant industries are then only very rarely affected (Albert et al. 2000). If they are, however, the effect is disproportionately larger due to cascading failures. By exploiting this asymmetry, the high-dimensional space of macroprudential oversight can thus be reduced to a much more manageable three percent of European industries.

Even though we conscientiously tested the robustness of all our results, our approach faces several limitations, essentially due to data constraints. Within the model, we completely abstracted from the monetary side of the economy, whereas both macroeconomic theory and our empirical evidence strongly hint at the presence and relevance of monetary channels (Ozdagli and Weber 2017). A possible remedy could be to introduce nominal frictions into the model. Empirically, the WIOD provides only rather short time series, limiting both the generalizability of our results and the number of controls we can reasonably include within the regression specification. Also, while our industry level model is an attempt at disaggregation, industries are of course themselves aggregate abstractions which average over economic decisions conducted at the firm or even plant level (Bernard et al. 2022). We are convinced that datasets without these caveats will pave numerous avenues for further research, not least to establish proper microfoundations for the relationships we only posit in reduced form at the moment. After all, our results underline the importance of mesoeconomic objects such as size distributions and contagion networks for proper aggregation from the micro to the macro level already in our present exposition as reduced form relationships.

## 2.A Appendix: Generic Residuals and Simple Linear Regression

Consider a generic weight  $w_i$  for industry  $i$ . For our two specifications, this weight could be the Domar weight for the granular case,  $S_i/Y$  and  $S_i/Y \cdot \theta_i$  for the mesoeconomic case. We assume that the TFP growth rate of the total economy is a weighted sum of

idiosyncratic productivity shocks to  $N$  industries which implies

$$\frac{dTFP_t}{TFP_t} = \sum_{i=1}^N w_{i,t} d\pi_{i,t},$$

where  $\pi_{i,t}$  is the productivity of firm  $i$  in period  $t$ ,  $w_{i,t}$  its weight at  $t$  and  $TFP_t$  is the Total Factor Productivity in period  $t$ . We assume that the idiosyncratic productivity shocks  $d\pi_i$  are independent of the weights  $w_i$  which, for the granular case, would imply Gibrat's law. In the following, we call this condition *Generalized Gibrat's Law*. Assume that that productivity shocks are Hicks-neutral and that GDP growth  $y_t$  is proportional to TFP growth. For these assumptions, the cumulative idiosyncratic shocks multiplied by a factor  $\psi$  yield the GDP growth rate  $y_t$  which is given by

$$y_t = \psi \cdot \sum_{i=1}^N w_{i,t} d\pi_{i,t}.$$

It follows that the variance of GDP growth rates is given by

$$\text{var}(y) = \psi^2 \cdot \sigma_\pi^2 \cdot H(N),$$

where  $\sigma_\pi^2$  is the variance of productivity growth rates and  $H$  the Hirschman-Herfindahl index (HHI) of the weights  $w$ . For generic weights  $w$ ,  $H(N)$  is therefore just the sum of  $N$  squared weights, that is,

$$H(N) = \sum_{i=1}^N (w_i)^2.$$

Consider now a generic residual  $\mathcal{X}_t$  based purely on observables given by

$$\mathcal{X}_t = \sum_{i=1}^K w_{i,t} d\pi_{i,t},$$

where  $K$  denotes the  $K$  largest industry weights and assuming that we identified the correct productivity differences  $d\pi_{i,t}$  for all  $K$  firms. If we assume the Generalized Gibrat's law to hold, that is, productivity growth rates to be independent of weights,  $\sigma_\pi$  has to be independent of  $H(K)$  as it is defined by the weight distribution of industries. Also, this implies that  $E[\sigma_\pi(K)] = \sigma_\pi(N)$ , that is, the variance of productivity growth rates of any subsample of  $K$  industries and  $K > 1$  is expectationally equivalent to the variance of productivity growth rates for the whole sample. The variance of generic residual  $\mathcal{X}$  for the  $K$  largest industries in terms of weights is given by

$$\text{var}(\mathcal{X}^K) = \sigma_\pi^2(K) \cdot H(K).$$

Taking expectations to define the expected induced variance yields

$$\begin{aligned} E[\text{var}(\mathcal{X}^K)] &= E[\sigma_\pi^2(K)] \cdot H(K) \\ &= \sigma_\pi^2 \cdot H(K), \end{aligned}$$

as  $H(K)$  is independent of  $\sigma_\pi$  by the generalized version of Gibrat's law.

Consider a simple linear regression with

$$y_t = \beta_0 + \beta_1 \cdot \mathcal{X}_t + \epsilon_t.$$

An OLS estimation would yield for this specification the estimated  $\hat{\beta}_1$

$$\hat{\beta}_1 = \frac{\text{cov}(y, \mathcal{X})}{\text{var}(\mathcal{X})}.$$

Rewriting the initial specification gives

$$y_t = \psi \cdot \mathcal{X}_t^N = \psi \cdot (\mathcal{X}_t^K + \mathcal{X}_t^{N-K}),$$

for all considered periods  $t$ , where  $\mathcal{X}^N$  denotes the generic residual for all  $N$  entities,  $\mathcal{X}^K$  for the largest  $K$  ones and  $\mathcal{X}^{N-K}$  for the remaining smallest  $N - K$  industries or firms. From this definition,  $\hat{\beta}_1$  can be expressed as

$$\hat{\beta}_1 = \frac{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K) \cdot (y_t - \bar{y})}{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K)^2}.$$

From  $\psi \cdot (\mathcal{X}_t^K + \mathcal{X}_t^{N-K}) = y_t$  and therefore  $\psi \cdot (\bar{\mathcal{X}}^K + \bar{\mathcal{X}}^{N-K}) = \bar{y}$ , we get

$$= \frac{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K) \cdot \psi \cdot ((\mathcal{X}_t^K + \mathcal{X}_t^{N-K}) - (\bar{\mathcal{X}}^K + \bar{\mathcal{X}}^{N-K}))}{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K)^2}.$$

Rearranging gives

$$\begin{aligned} &= \psi \cdot \frac{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K) \cdot ((\mathcal{X}_t^K - \bar{\mathcal{X}}^K) + (\mathcal{X}_t^{N-K} - \bar{\mathcal{X}}^{N-K}))}{\sum_t^T (\mathcal{X}_t^K - \bar{\mathcal{X}}^K)^2} \\ &= \psi \cdot \frac{\text{var}(\mathcal{X}^K) + \text{cov}(\mathcal{X}^K, \mathcal{X}^{N-K})}{\text{var}(\mathcal{X}^K)} \\ &= \psi \cdot \left(1 + \frac{\text{cov}(\mathcal{X}^K, \mathcal{X}^{N-K})}{\text{var}(\mathcal{X}^K)}\right). \end{aligned}$$

Since we assume that all sectors draw independently from productivity distributions with equal first and second moments,  $\text{cov}(\mathcal{X}^K, \mathcal{X}^{N-K})$  and  $\text{var}(\mathcal{X}^K)$  are independent. Also, by the same independence assumption, it has to hold that  $E[\text{cov}(\mathcal{X}^K, \mathcal{X}^{N-K})] = 0$ . Taking

expectations yields

$$\begin{aligned} E[\hat{\beta}_1] &= E[\psi \cdot (1 + \frac{cov(\mathcal{X}^K, \mathcal{X}^{N-K})}{var(\mathcal{X}^K)})] \\ &= \psi + \psi \cdot E[\frac{cov(\mathcal{X}^K, \mathcal{X}^{N-K})}{var(\mathcal{X}^K)}] \end{aligned}$$

and by invoking the independence of  $cov(\mathcal{X}^K, \mathcal{X}^{N-K})$  and  $var(\mathcal{X}^K)$

$$= \psi + \psi \cdot (E[cov(\mathcal{X}^K, \mathcal{X}^{N-K})] \cdot E[var(\mathcal{X}^K)^{-1}])$$

which yields with  $E[cov(\mathcal{X}^K, \mathcal{X}^{N-K})] = 0$  again from independence

$$= \psi.$$

Consider now the coefficient of determination for the given simple linear regression with

$$R^2 = \hat{\beta}_1^2 \cdot \frac{var(\mathcal{X}^K)}{var(y)}.$$

Taking expectations yields

$$E[R^2] = E[\hat{\beta}_1^2 \cdot \frac{var(\mathcal{X}^K)}{var(y)}]$$

and since  $var(y)$  is observable,  $E[var(y)] = var(y)$  and by invoking that  $\hat{\beta}_1^2$  and  $var(\mathcal{X}^K)$  are uncorrelated,<sup>24</sup>

$$= var(y)^{-1} \cdot E[\hat{\beta}_1^2] \cdot E[var(\mathcal{X}^K)].$$

Substituting gives

$$\begin{aligned} &= \frac{\psi^2 \cdot \sigma_\pi^2 \cdot H(K)}{\psi^2 \cdot \sigma_\pi^2 \cdot H(N)} \\ &= \frac{H(K)}{H(N)} \\ &= \sum_{i=1}^K \frac{(w_i)^2}{\sum_{j=1}^N (w_j)^2}. \end{aligned}$$

---

<sup>24</sup>For  $\hat{\beta}_1^2$  and  $var(\mathcal{X}^K)$  to be uncorrelated, we need to show that the correlation of  $var(\mathcal{X}^K)$  and  $cov(\mathcal{X}^K, \mathcal{X}^{N-K})/var(\mathcal{X}^K)$  is zero. Since  $var(\mathcal{X}^K)$  and  $cov(\mathcal{X}^K, \mathcal{X}^{N-K})$  are independent,  $cov(cov(\mathcal{X}^K, \mathcal{X}^{N-K})/var(\mathcal{X}^K), var(\mathcal{X}^K)) = E[cov(\mathcal{X}^K, \mathcal{X}^{N-K})] \cdot (1 - E[var(\mathcal{X}^K)] \cdot E[1/var(\mathcal{X}^K)])$  and by  $E[cov(\mathcal{X}^K, \mathcal{X}^{N-K})] = 0$ , we thus get  $cov(cov(\mathcal{X}^K, \mathcal{X}^{N-K})/var(\mathcal{X}^K), var(\mathcal{X}^K)) = 0$ , i.e.,  $\hat{\beta}_1^2$  and  $var(\mathcal{X}^K)$  being uncorrelated.

Thus, under the generalized version of Gibrat's law and assuming that we correctly identified the idiosyncratic productivity shocks to each firm or industry, the expected  $R^2$  for the granular residual for  $K$  industries should be given by the ratio of the weight Herfindahl for the  $K$  largest firms or industries to the weight Herfindahl for the total economy. To illustrate this result, we perform a simple Monte Carlo simulation. The weight distribution is given by a power-law with tail exponent  $\alpha = 1.9$ . There are  $N = 100$  industries in total and the idiosyncratic productivity shocks are drawn from a Laplacian distribution with location parameter  $\mu = 0$  and scale parameter  $\sigma = 0.05$ . We consider 100 time periods each and repeat the simulation 50 times with different realizations for the idiosyncratic shocks. Figure 2.3 shows the mean  $R^2$  for these 50 simulation runs and a simple linear regression model as presented above as well as the expected  $R^2$  by the results below. This is given by  $\frac{H(K)}{H(N)} = \frac{\zeta_K(\frac{2}{\alpha})}{\zeta_N(\frac{2}{\alpha})}$  where  $\zeta_N(\alpha)$  is a truncated Zeta function at  $N$  with  $\zeta_N(\alpha) = \sum_{k=1}^N k^{-\alpha}$  and with  $\alpha = 1.9$ .

As can easily be seen, the expected  $R^2$  from the ratio of Herfindahls is remarkably close to the mean  $R^2$  from the simulation exercise. Notice that our results are not materially sensitive to the specific choice of the distribution for idiosyncratic productivity shocks. The results are robust for several different parameter values of  $\mu$ ,  $\sigma$  and  $\kappa$  for a (symmetrical) Subbotin distribution that includes the outlined Laplacian as a special case for  $\kappa = 1$ . The approximation only breaks down for very small shape parameters  $\kappa < 0.3$ , for which the variance and all higher moments explode such that 100 different realizations of idiosyncratic productivity shocks no longer warrant convergence of the mean to the expected value by the law of large numbers. The result is also robust to various specifications of the weight distribution (measured by the tail exponent  $\alpha$ ).<sup>25</sup> Of course, this is only the expected or mean functional form of the  $R^2$ . Indeed, depending on the specific realization of draws from the distribution for idiosyncratic shocks, the realized  $R^2$  might differ, as Figure 2.4 shows. Given a sufficiently heterogeneous weight distribution, however, the  $R^2$  plot cannot deviate too much from the expected  $R^2$ , as shown in Figure 2.3, since it should increase almost monotonically, thereby excluding large downturns in the empirical  $R^2$  plot.

Finally, consider the boundary condition Gabaix (2011) gives for granular firms to exist - they have to be distributed by a power law with a tail exponent  $\alpha$  at most equal to 2. This implies that if all  $K$  weights are granular,  $H(K) \geq \zeta_K(1)$ . This, however, is just the harmonic series and thus, for any given  $H(N) = H$  for the total economy, the ratio  $\frac{H(K)}{H} \approx \frac{1}{H(N)}(\ln(K) + \gamma)$ , where  $\gamma$  is the Euler-Mascheroni constant. Thus the  $R^2$  plot can be expected to increase at least in proportion to the natural logarithm of  $K$ .

---

<sup>25</sup> Material available upon request.



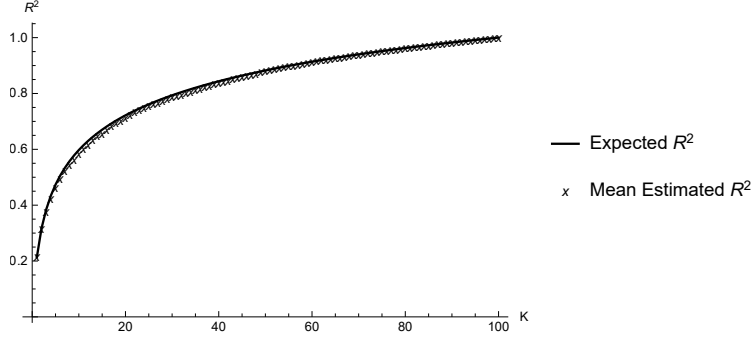


Figure 2.3: Mean  $R^2$  for Monte Carlo simulation exercise and expected  $R^2$  derived from the weight distribution.

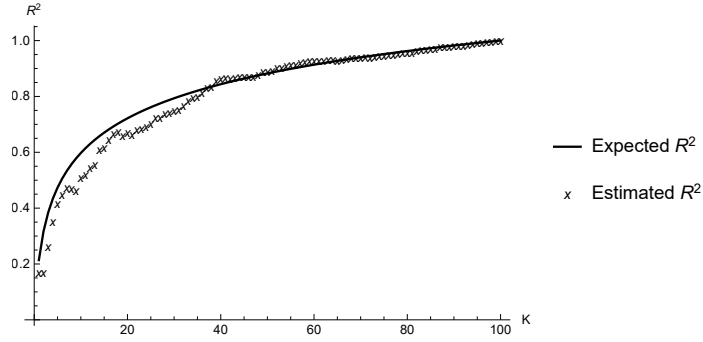


Figure 2.4:  $R^2$  for a specific realization of idiosyncratic productivity shocks and expected  $R^2$  derived from the weight distribution.

## 2.B Appendix: Winsorizing, $\beta$ estimates and $R^2$

Consider the same framework as in Appendix 2.A. However, assume now that the productivity growth rate distribution of the  $K$  largest industries was winsorized at the  $p$ th and  $1 - p$ th quantile. We define  $\nu(p) = \tilde{\sigma}_\pi^2(K, p) / \sigma_\pi^2(K)$ , that is,  $\nu$  gives the ratio of variances for the  $p$ -winsorized and initial distribution of productivity growth rates for the  $K$  largest industries with  $\nu(p) \in (0, 1)$  for  $p \in (0, 0.5)$ . It has to hold that  $\lim_{p \rightarrow 0} \nu(p) = 1$  and  $\lim_{p \rightarrow 0.5} \nu(p) = 0$ . If we assume the generalized version of Gibrat's law to hold, it follows that  $E[\sigma_\pi^2(K)] = E[\sigma_\pi^2(N)] = \sigma_\pi^2$  and therefore  $E[\tilde{\sigma}_\pi^2(K, p)] = \nu(p) \cdot \sigma_\pi^2$ .

The expected, winsorized variance of  $\tilde{\mathcal{X}}^K(p)$  of the generic residual defined in the above section is then

$$E[\text{var}(\tilde{\mathcal{X}}^K(p))] = \nu(p) \cdot \sigma_\pi^2 \cdot H(K).$$

For analytical tractability, we assume that the winsorized generic residual is affected by some time-invariant scalar  $\delta^K(p) \in \mathbb{R}$  such that  $\mathcal{X}_t^K = \tilde{\mathcal{X}}_t^K(p) + \delta^K(p) \forall t$ , which might vary with the winsorizing probability  $p$  and with the number of considered entities  $K$ .

Following the discussion in Appendix 2.A, we thus get

$$\begin{aligned} y_t &= \psi \cdot \mathcal{X}_t^N \\ &= \psi \cdot (\mathcal{X}_t^K + \mathcal{X}_t^{N-K}) \\ &= \psi \cdot (\tilde{\mathcal{X}}_t^K(p) + \delta^K(p) + \mathcal{X}_t^{N-K}) \end{aligned}$$

and analogously for the time-series averages, where  $\bar{\mathcal{X}}^K(p)$  is the time-series average of the generic residual winsorized at  $p$

$$\bar{y} = \psi \cdot (\bar{\mathcal{X}}^K(p) + \delta^K(p) + \bar{\mathcal{X}}^{N-K}).$$

For the regression coefficient, it follows that

$$\begin{aligned} \hat{\beta}_1(p) &= \frac{\sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p)) \cdot (y_t - \bar{y})}{\sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p))^2} \\ &= \left[ \sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p)) \cdot \psi \cdot ((\tilde{\mathcal{X}}_t^K(p) + \delta^K(p) + \mathcal{X}_t^{N-K}) - (\bar{\mathcal{X}}^K(p) + \delta^K(p) + \bar{\mathcal{X}}^{N-K})) \right] / \left[ \sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p))^2 \right] \\ &= \psi \cdot \frac{\sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p)) \cdot (\tilde{\mathcal{X}}_t^K(p) + \mathcal{X}_t^{N-K} - \bar{\mathcal{X}}^K(p) + \bar{\mathcal{X}}^{N-K})}{\sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p))^2} \end{aligned}$$

and by rearranging

$$\begin{aligned} &= \psi \cdot \frac{\sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p)) \cdot (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p)) + (\mathcal{X}_t^{N-K} - \bar{\mathcal{X}}^{N-K})}{\sum_t^T (\tilde{\mathcal{X}}_t^K(p) - \bar{\mathcal{X}}^K(p))^2} \\ &= \psi \cdot \frac{\text{var}(\tilde{\mathcal{X}}^K(p)) + \text{cov}(\tilde{\mathcal{X}}^K(p), \mathcal{X}^{N-K})}{\text{var}(\tilde{\mathcal{X}}^K(p))} \\ &= \psi \cdot \left( 1 + \frac{\text{cov}(\tilde{\mathcal{X}}^K(p), \mathcal{X}^{N-K})}{\text{var}(\tilde{\mathcal{X}}^K(p))} \right). \end{aligned}$$

If the winsorizing procedure did not generate spurious dependence between industries, the terms  $\text{cov}(\tilde{\mathcal{X}}^K(p), \mathcal{X}^{N-K})$  and  $\text{var}(\tilde{\mathcal{X}}^K(p))$  remain independent and  $E[\text{cov}(\tilde{\mathcal{X}}^K(p), \mathcal{X}^{N-K})] = 0$ . Thus, as shown in Appendix 2.A, the estimate for the  $\beta_1$  coefficient in a simple linear regression model remains unbiased, and  $E[\hat{\beta}_1(K)|p = 0] = E[\hat{\beta}_1(K)|p] = \psi$ . However, the expected  $R^2$  as the explanatory power for  $K$  industries should for any degree of winsorizing  $p$  be biased downwards. This follows purely from the fact that the variance of the winsorized generic residual is downwards biased in expecta-

tions. To see this, consider  $E[R^2(K)|p]$  as

$$\begin{aligned} E[R^2(K)|p] &= \psi^2 \frac{E[\text{var}(\tilde{\mathcal{X}}^K(p))]}{\text{var}(y)} \\ &= \frac{\psi^2 \cdot \nu(p) \cdot \sigma_\pi^2 \cdot H(K)}{\psi^2 \cdot \sigma_\pi^2 \cdot H(N)} \\ &= \nu(p) \cdot \frac{H(K)}{H(N)} \\ &= \nu(p) \cdot \sum_{i=1}^K \frac{(w_i)^2}{\sum_{j=1}^N (w_j)^2} \end{aligned}$$

and that  $\nu(p) < 1$  for all  $p \in (0, 0.5)$ . It follows that, compared to the benchmark expected  $R^2$  without winsorizing, the estimate is downwards biased by  $\Lambda(p)$  percent given by

$$\begin{aligned} \Lambda(p) &= \frac{E[R^2(K)|p=0] - E[R^2(K)|p]}{E[R^2(K)|p=0]} \\ &= 1 - \nu(p). \end{aligned}$$

The degree of bias crucially depends on the kurtosis of the underlying growth rate distribution, as we illustrate in Appendix C available at [https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi\\_lehrstuehle/vwl\\_internationale\\_wirtschaft/Online\\_Appendix\\_SGM.pdf](https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/sowi_lehrstuehle/vwl_internationale_wirtschaft/Online_Appendix_SGM.pdf).

## References

- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016.
- Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382.
- Ascari, G., Fagiolo, G., and Roventini, A. (2015). Fat-tail distributions and business cycle models. *Macroeconomic Dynamics*, 19(2):465–476.
- Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal: Macroeconomics*, 9(4):254–80.
- Baqae, D. and Rubbo, E. (2022). Micro propagation and macro aggregation. *Annual Review of Economics*, forthcoming.
- Baqae, D. R. and Farhi, E. (2019). The macroeconomic impact of microeconomic shocks: Beyond Hulten’s theorem. *Econometrica*, 87(4):1155–1203.
- Barro, R. J. (1999). Notes on growth accounting. *Journal of Economic Growth*, 4(2):119–137.

- Barrot, J.-N. and Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *Quarterly Journal of Economics*, 131(3):1543–1592.
- Bernard, A. B., Dhyne, E., Magerman, G., Manova, K., and Moxnes, A. (2022). The origins of firm heterogeneity: A production network approach. *Journal of Political Economy*, 130(7):1765 – 1804.
- Blanco-Arroyo, O., Ruiz-Bufo, A., Vidal-Tomás, D., and Alfarano, S. (2018). On the determination of the granular size of the economy. *Economics Letters*, 173:35–38.
- Boehm, C. E., Flaaen, A., and Pandalai-Nayar, N. (2019). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 Tōhoku earthquake. *Review of Economics and Statistics*, 101(1):60–75.
- Bonacich, P. (1991). Simultaneous group and individual centralities. *Social Networks*, 13(2):155–168.
- Bottazzi, G., Li, L., and Secchi, A. (2019). Aggregate fluctuations and the distribution of firm growth rates. *Industrial and Corporate Change*, 28(3):635–656.
- Carvalho, V. M. (2014). From micro to macro via production networks. *Journal of Economic Perspectives*, 28(4):23–47.
- Carvalho, V. M. and Tahbaz-Salehi, A. (2019). Production networks: A primer. *Annual Review of Economics*, 11:635–663.
- Daniele, F. and Stüber, H. (2020). The micro-origins of business cycles: Evidence from German metropolitan areas. *Review of Economics and Statistics*, forthcoming.
- De Grauwe, P. (2020). *Economics of the monetary union*. Oxford University Press.
- Di Giovanni, J., Levchenko, A. A., and Méjean, I. (2014). Firms, destinations, and aggregate fluctuations. *Econometrica*, 82(4):1303–1340.
- Domar, E. D. (1961). On the measurement of technological change. *Economic Journal*, 71(284):709–729.
- Dosi, G., Grazzi, M., Tomasi, C., and Zeli, A. (2012). Turbulence underneath the Big Calm? The micro-evidence behind Italian productivity dynamics. *Small Business Economics*, 39(4):1043–1067.
- Dosi, G., Napoletano, M., Roventini, A., and Treibich, T. (2019). Debunking the granular origins of aggregate fluctuations: From real business cycles back to Keynes. *Journal of Evolutionary Economics*, 29(1):67–90.
- Ebeke, M. C. H. and Eklou, K. M. (2017). The granular origins of macroeconomic fluctuations in Europe. *IMF Working Paper Series*, 17(229).
- Foerster, A. T., Sarte, P.-D. G., and Watson, M. W. (2011). Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy*, 119(1):1–38.

- Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772.
- Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives*, 30(1):185–206.
- Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.
- Gorman, W. M. (1961). On a class of preference fields. *Metroeconomica*, 13(2):53–56.
- Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26):10570–10575.
- Hulten, C. R. (1978). Growth accounting with intermediate inputs. *Review of Economic Studies*, 45(3):511–518.
- Imry, Y. (2008). *Introduction to mesoscopic physics*. Oxford University Press.
- Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43.
- Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives*, 6(2):117–136.
- Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods & Research*, 33(2):188–229.
- Liu, E. (2020). Industrial policies in production networks. *Quarterly Journal of Economics*, 134(3):1883–1948.
- Lucas, R. E. (1977). Understanding business cycles. In *Carnegie-Rochester Conference Series on Public Policy*, volume 5, pages 7–29. Elsevier.
- Miranda-Pinto, J. and Shen, Y. (2019). A granular view of the Australian business cycle. *Economic Record*, 95(311):407–424.
- Ozdagli, A. and Weber, M. (2017). Monetary policy through production networks: Evidence from the stock market. *NBER Working Paper Series*, 23424.
- Saint-Paul, G. (1993). Productivity growth and the structure of the business cycle. *European Economic Review*, 37(4):861–883.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1):1–48.
- Snowdon, B., Vane, H. R., and Wynarczyk, P. (1994). *A modern guide to macroeconomics: An introduction to competing schools of thought*. Edgar Elgar Publishing.
- Stella, A. (2015). Firm dynamics and the origins of aggregate fluctuations. *Journal of Economic Dynamics and Control*, 55:71–88.
- Stumpf, M. P. and Porter, M. A. (2012). Critical truths about power laws. *Science*, 335(6069):665–666.

- Summers, L. H. (1986). Some skeptical observations on real business cycle theory. *Federal Reserve Bank of Minneapolis Quarterly Review*, Fall:23–27.
- Sutton, J. (1997). Gibrat’s legacy. *Journal of Economic Literature*, 19:40–59.
- Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy*, 112(6):1181–1222.
- Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., and Vries, G. J. (2015). An illustrated user guide to the world input-output database: The case of global automotive production. *Review of International Economics*, 23(3):575–605.
- Wagner, J. and Weche, J. P. (2020). On the granularity of the German economy—first evidence from the top 100 companies panel database. *Applied Economics Letters*, 27(21):1768–1771.
- World Bank (2018). The World Bank data. <https://data.worldbank.org/indicator/NY.GDP.PCAP.KD.ZG?locations=SE-FI-EU>. [Online; accessed 2018-07-04].
- World input-output database (2017). WIOD home. <http://www.wiod.org/home>. [Online; accessed 2017-04-13].
- Yu, X., Dosi, G., Lei, J., and Nuvolari, A. (2015). Institutional change and productivity growth in China’s manufacturing: The microeconomics of knowledge accumulation and “creative restructuring”. *Industrial and Corporate Change*, 24(3):565–602.

# How Wealthy Are the Rich?

---

Joint work with M. Milaković.

Published under Creative Commons Attribution 4.0 License:

Schulz, J. and Milaković, M. (2023). How wealthy are the rich?. *Review of Income and Wealth*, 69(1):100–123. <https://doi.org/10.1111/roiw.12550>

We thank Simone Alfarano, Anna Gebhard, Philipp Mundt, our colleagues at the Bamberg Research Group on Behavioral Macroeconomics, as well as the participants of the 1st Behavioral Macroeconomics Workshop in Bamberg, the 12th International Conference on Computational and Financial Econometrics in Pisa, the 23rd Forum for Macroeconomics and Macroeconomic Policies, the 9th Young Economists Conference and the 77th Annual Congress of the International Institute of Public Finance for helpful comments and suggestions at crucial junctions of this investigation. Financial support by the Hans-Böckler-Foundation through grant PK045 is gratefully acknowledged.

---

## Abstract

Underreporting and undersampling biases in top tail wealth, although widely acknowledged, have not been statistically quantified so far, essentially because they are not readily observable. Here we exploit the functional form of power law-like regimes in top tail wealth to derive analytical expressions for these biases, and employ German microdata from a popular survey and rich list to illustrate that tiny differences in non-response rates lead to tail wealth estimates that differ by an order of magnitude, in our case ranging from one to nine trillion euros. Underreporting seriously compounds the problem, and we find that the estimation of totals in scale-free systems oftentimes tends to be spurious. Our findings also suggest that recent debates on the existence of scale- or type-dependence in returns to wealth are ill-posed because the available data cannot discriminate between scale- or type-dependence on the one hand, and statistical biases on the other. Yet both economic theory and mathematical formalism indicate that sampling and reporting biases are more plausible explanations for the observed data than scale- or type-dependence.

---

---

## Contents of this Paper

3.1	Introduction . . . . .	49
3.2	Model . . . . .	53
3.3	Data and Estimation . . . . .	55
3.3.1	Data . . . . .	56
3.3.2	Estimation . . . . .	56
3.4	Results . . . . .	58
3.4.1	Distributional Results . . . . .	59
3.4.2	Data First . . . . .	60
3.4.3	Theory First . . . . .	62
3.4.4	Total Wealth Estimates . . . . .	65
3.5	Discussion . . . . .	67
3.A	Model . . . . .	69
3.B	Analytical Results for the Estimation of $\alpha$ under $q$ -Truncation . . . . .	71
3.C	Empirical Densities of Returns to Wealth Portfolios . . . . .	75
3.D	Total Wealth Estimates . . . . .	76
3.E	Equiprobable Sampling From a Power Law . . . . .	77
3.F	Logarithmic Sampling From a Power Law . . . . .	78
	References . . . . .	79

---



### 3.1 Introduction

The starting point for this analysis was a conscientious effort to quantify the total wealth of the richest Germans from survey microdata. This seemingly innocuous exercise pointed us to a problem which, to the best of our knowledge, has not yet been adequately addressed in the pertinent literature. The problem arises in the top tail of wealth, generally following power law-like distributions, where survey data apparently suggest total wealth to be orders of magnitude smaller than implied by named rankings of the super-rich, often referred to as rich lists. Extrapolating the power law *backward* from observed top wealth levels to some unobserved minimum is asymptotically unbiased. Severe biases can arise, however, when extrapolating *forward* from relatively low levels to unreliable or missing maximum wealth levels (Cristelli et al. 2012). In survey data the latter typically leads to strongly downward biased estimates of wealth and inequality (Eckerstorfer et al. 2016; Vermeulen 2018). Since we cannot quantify this effect without data that go beyond the available, we propose two limit interpretations to gauge the potential impact of this bias. In what we term the *data first* limit, we assume both upper and lower truncated samples to deliver unbiased estimates. Put differently, we attribute all observed differences between upper and lower truncated samples to truly existing differences in the data generating process. In the complementary *theory first* limit, we assume the data generating process to be homogeneous across samples on the different scales, attributing the entire observed difference to statistical bias. We show that tail wealth estimates differ by an order of magnitude, depending on which of the two pre-analytical visions we employ.

The literature so far has implicitly taken a *data first* stance on this issue (Eckerstorfer et al. 2016; Vermeulen 2016, 2018; Bach et al. 2019). Our primary goal with this paper is to argue that a *theory first* perspective is at the very least equally plausible. To show this we introduce different categories of biases that affect measured inequality, and provide closed-form expressions that are readily estimated. First, we show that *underreporting* incentives by themselves are insufficient to generate biased estimates, as the estimate is asymptotically unbiased if the entire population unanimously underreports their wealth. Inequality is underestimated only if underreporting is more pronounced for the richest, which seems intuitively plausible as the super-rich have mightier means at their disposal to avoid taxes than the average person or household. Second, we demonstrate that differential underreporting by the super-rich indeed leads to downward biased estimates of inequality for the entire population. Finally, and most importantly, the impact of underreporting rates is highly non-linear. Even if only a fraction of actual wealth is reported, this will greatly reduce the resulting bias compared to when information on a fraction of the super-rich is missing altogether. We call the latter case *undersampling*, which is typical of survey data that essentially use equiprobable sampling and therefore do not adequately capture the richest individuals in power law-like regimes. We also show that

*logarithmic sampling* would greatly improve the statistical quality of wealth surveys. The named rich lists, on the other hand, will be subject to reporting biases as they explicitly try to account for the super-rich but typically suffer from data availability and salience issues, as well as adverse (tax) reporting incentives. Without additional information, both the estimated underreporting and undersampling rates remain within plausible bounds, so the polar *data first* and *theory first* perspectives would appear equally plausible at first. While it is hardly surprising that the two perspectives imply different estimates for top tail wealth, the difference turns out to be enormous. The lowest estimate arising from *data first* is around one trillion euros for Germany's top tail wealth, while the *theory first* estimates reach about nine trillion euros. These vast differences, spanning almost one order of magnitude in top tail wealth, are caused by tiny non-response rates on the order of a tenth of a percent. This disconcerting result suggests that aggregate findings within the *data first* framework can become heavily distorted by tiny degrees of undersampling. The severity of the problem extends far beyond the German dataset since our results are functions of the power law tail of wealth distributions that applies across many countries and time periods. Consequently, estimates of total wealth will crucially depend on the pre-analytical perspective and should thus be treated with extreme caution. If total wealth estimates are to be stated, we believe that scientific integrity at least demands to report the range from the smallest estimates of a *data first* perspective to the largest estimates of a *theory first* perspective, especially if these estimates are intended to inform economic policy or public debate.

The ubiquity of power laws has led to numerous suggestions for potential generating mechanisms, reviewed for instance by Gabaix (2009) or Luttmner (2010). In the case of top tail wealth, any candidate mechanism should be based on a property that is common across the various time periods, countries, or proxies of wealth. One common property, at least across the different varieties of capitalism, concerns the primary types of assets in super-rich portfolios, namely entrepreneurial stakes, financial assets, and speculative (that is non-owner occupied) real-estate, which are perpetually reinvested into or reallocated among these asset classes (Davies and Shorrocks 2000; Wachter and Yogo 2010).<sup>1</sup> Thus a random growth model featuring a multiplicative component seems to be the most adequate candidate for a sensible generating mechanism. The idea to explain the emergence of power law tails with stochastic multiplicative processes has a long history but has fallen out of fashion in economics for many decades, essentially for its lack of microfoundations.

---

<sup>1</sup>From an accounting standpoint, this perpetual reallocation and investment is closely related to saving and there is a consensus in the literature that propensities to save are strongly positively correlated with (lifetime) income or wealth (Dynan et al. 2004; Jappelli and Pistaferri 2014). This also holds for entrepreneurial households (Quadrini 1999). As a major reason for this relationship, Deaton (2003) identifies credit constraints that are only binding for low wealth households and individuals. Alan et al. (2015), on the other hand, provide a critical discussion of the identification strategy and find no differential savings behavior with respect to long-term income. The major limitation of Alan et al. (2015) is the exclusion of the wealthiest one percent that we are primarily concerned with here.

Yet random multiplicative growth has recently regained traction within economically motivated partial and general equilibrium models that endogenously generate power law tails in wealth from stochastic capital or asset accumulation (Levy 2003; Levy and Levy 2003; Nirei and Souma 2007; Nirei 2009; Benhabib et al. 2011; Toda 2014; Piketty and Zucman 2015; Hubmer et al. 2016; Aoki and Nirei 2016; Benhabib and Bisin 2018).

The literature on random multiplicative growth has typically placed weak restrictions on the particular form of return distributions governing the stochastic process. One notable exception, however, is the assumption of an equilibrating tendency for the expected (risk-adjusted) rate of return or, in more technical terms, of a homogeneous return distribution across wealth portfolios. This is consistent with the classical notion of competition, the implications of (semi-strong) informationally efficient capital markets, and the idea that investors' superior talent in either fundamental or technical analysis cannot lead to excess returns over extended periods of time (Fama 1965, 1970, 1991). Indeed, as Levy (2003) and Levy and Levy (2003) show both experimentally and via Monte Carlo simulations, the scope for differential talent is very limited in light of power law distributed top wealth. If one group of investors were to consistently outperform another group of less talented investors in terms of their expected returns by only a tiny margin, the functional form of the emergent stationary distribution would differ significantly from a power law and exhibit concavity on double-logarithmic scale.<sup>2</sup> Hence the defining characteristic of the *theory first* perspective is to assume a homogeneous return distribution, thereby implying equivalent data generating processes across samples.

A more recent strand of literature has started to challenge the homogeneity hypothesis on both theoretical and empirical grounds. Bach et al. (2017) and Fagereng et al. (2020) find excess risk-adjusted returns for the wealthiest portfolios, the latter even claiming persistence in abnormal returns, indicating persistent heterogeneity in financial information and talent *if we take the data at face value*. From a more theoretical perspective, Luttmner (2011) and Gabaix et al. (2016) build on the well-known limitation of random growth models to typically generate very slow transitions. The former puts this in terms of the stationary distribution of assets, with a half-life of assets that would be way too high from an empirical point of view, while the latter argue (formally equivalently) that the rate of convergence to the new stationary distribution after a shock to the variance in the permanent component of earnings is too slow to account for the observed rapid rise in top-level income inequality. Gabaix et al. (2016) and Jones and Kim (2018) thus put forward the hypothesis of heterogeneous returns to explain the observed rise in income and wealth inequality, whereby excess returns are either correlated with wealth levels ("scale-dependence") or result from differential talent ("type-dependence"). In in-

---

<sup>2</sup>They consider two Gaussian return distributions that merely differ in expected value, showing that already a difference by one percentage point in expected returns leads to a stationary distribution that significantly differs from the Pareto type.

formationally efficient capital markets, scale-dependence can only occur when the set of investment opportunities increases in wealth. Hedge funds and some private banks perhaps provide anecdotal evidence, as hedge funds typically require high minimum investment inlays (King and Maier 2010), while some private banks like JP Morgan Chase require their private clients to hold at least ten million dollars (Glazer 2016). Concerning type-dependence, Gabaix et al. (2016) circumvent the formal problem that differential talent is inconsistent with a Pareto distribution by essentially assuming that “high growth types” only stay in the high growth regime for a limited amount of time and cannot return there. This idea not only lacks theoretical appeal, it also introduces another degree of freedom into any empirical investigation that now has to justify after how many periods of abnormally high returns one can safely claim type-dependence.

Moreover, given that our data lack information on investors’ sophistication, this notion of type-dependence is phenomenologically equivalent to scale-dependence since we cannot control for investors’ ability. Put differently, we cannot distinguish between the hypothesis that individuals are rich because of their excess returns, and the alternative hypothesis that they have excess returns because they are rich. We will thus only focus on testing for scale-dependence. This hypothesis corresponds to the *data first* interpretation, as observed differences between high and low scale samples are then assumed to reflect true differences in the data generating process, that is to say scale-dependent random growth. We will argue, however, that the idea of scale-dependent growth is not only problematic from a formal point of view, but that it also violates economic intuitions like informational efficiency or the classical concept of competition that predicts a tendency for the equalization of returns. *Theory first* leaves these economic intuitions intact by attributing observed deviations in the data to statistical biases arising from undersampling and underreporting, and also casts a different light on the apparently reversed risk-return trade-off that we observe in the data.

The remainder of this paper is organized as follows: section 3.2 derives the biases in estimates of the tail exponent that arise from underreporting and undersampling, respectively. Throughout the paper we have relegated all derivations to the appendix in order to emphasize important conceptual differences over technical detail. Section 3.3 introduces the data and discusses our estimation procedure. Our results are presented in section 3.4, where we put forward two mutually exclusive yet on their own reasonably plausible explanations for the observed behavior in the data. Section 3.5 discusses the implications of our results for existing work on top tail wealth, and concludes with the suggestion to improve future surveys through logarithmic sampling.

## 3.2 Model

The *data first* and *theory first* interpretations are purposefully designed to be antithetical, although we will show that the formal explication of both interpretations can be reduced to conceptually closely related mechanisms that affect measured tail inequality at different stages of empirical estimation. For both interpretations, Zipf's (1949) law with a tail exponent of unity is an attractor for a parsimonious stochastic multiplicative process that does not exhibit scale-dependence in accumulation or reporting. Consequently, an observed tail exponent that differs from unity implies scale-dependent behavior in both frameworks. *Data first* attributes this to scale-dependent stochastic growth at the level of accumulation, while *theory first* assumes that it is fully caused by scale-dependent reporting behavior at the level of measurement. Though the mechanisms are formally quite similar, the pre-analytical vision obviously differs substantially between narratives. Within the *data first* framework, the measured tail inequality is a sufficient statistic for both the true (snapshot) inequality among the richest and scale-dependence within the wealth accumulation process over time, because tail inequality is intricately linked to the nature of the underlying stochastic process of multiplicative growth. In Appendix 3.A , we consider a standard drift-diffusion process to show that the tail index of the stationary power law distribution is uniquely determined by the expected return and variance of the stochastic growth process. As Gabaix (1999) shows, and we rederive in greater detail in Appendix 3.A , the stationary distribution of the right tail for this type of general process is a power law, with tail index  $\alpha$  given by

$$\alpha(w; \gamma(\cdot), \sigma(\cdot)) = 1 - 2 \frac{\gamma(w) - \bar{\gamma}}{\sigma^2(w)} + \frac{w}{\sigma^2(w)} \frac{\partial \sigma^2(w)}{\partial w} \text{ with } w \in \mathbb{R}^+, \quad (3.1)$$

where  $\bar{\gamma}$  is the average wealth growth rate and  $\gamma(w)$  the (normalized) mean growth rate for a given wealth level  $w$ . Expression (3.1) has intuitive comparative statics with respect to the degree of scale-dependence in both mean growth rates  $\gamma$  and variance  $\sigma^2$ . Whenever the expected (excess) mean growth rate  $\gamma(w) - \bar{\gamma}$  increases in wealth, the tail exponent decreases and stationary inequality rises. Thus positive scale-dependence in expected returns increases measured inequality. When variance exhibits positive scale-dependence,  $\partial \sigma^2(w)/\partial w > 0$ , tail exponents increase and system-wide inequality hence decreases. Zipf's law with  $\alpha = 1$  is an interesting limit case for a situation without any scale-dependence (positive or negative), that is  $\gamma(w) = \bar{\gamma}, \forall w \in \mathbb{R}^+$ , and  $\partial \sigma^2(w)/\partial w = 0$ . These two conditions are typically called Gibrat's law after the seminal study by Gibrat (1931). Therefore Gibrat's law in growth rates is a sufficient condition for Zipf's law to hold in wealth levels. Córdoba (2008a,b) proves that it is also a necessary condition. The *data first* interpretation takes the Zipf benchmark in the stationary distribution as an

indication for scale-independence, while statistically significant deviations are evidence to the contrary.

The *data first* approach thus implicitly assumes that the measured tail inequality  $\hat{\alpha}$  is equivalent to the true stationary  $\alpha$  or, at least, that the estimate is not systematically biased in any direction. The polar *theory first* interpretation assumes no systematic scale-dependence of either type, that is  $\alpha = 1$ , and attributes significant deviations from this Zipf benchmark to underreporting and undersampling biases. While the relevance of distorted or missing observations has already been argued on empirical grounds and in Monte Carlo simulations (Vermeulen 2016, 2018), we derive closed-form expressions here that quantify the resulting bias in the tail exponent when the number of observations, denoted  $N$ , becomes large.<sup>3</sup> In addition, we will differentiate between unanimous and differential reporting behavior on the one hand, and underreporting versus undersampling on the other. Since undersampling or underreporting rates are impossible to estimate by the very nature of the problem, we consider three stylized scenarios that are analytically tractable: i) unanimous (proportional) underreporting, ii) differential (proportional) underreporting and iii) undersampling.

First, we consider the case of *unanimous underreporting*, that is all respondents only report a fraction  $\rho$  of their wealth. Call this fraction the *reporting rate*. We show in Appendix 3.B that this leads to an unbiased estimator of the tail index, hence unanimous underreporting does not pose problems for the estimation of inequality. This holds symmetrically for unanimous overreporting,  $\rho > 1$ , also showing that the estimator is invariant with respect to inflation. Whenever there are differential reporting rates, however, the bias is unambiguously positive and thus underestimates inequality. We call this case *differential underreporting*. For this, consider the case where the upper  $q$ -quantile of the wealth distribution only reports a fraction  $\rho$  of their wealth, from whence we show that for large  $N$  the estimated tail index, now denoted  $\hat{\alpha}^{du}$ , will differ from the Zipf benchmark such that

$$\hat{\alpha}^{du}(q, \rho) = \frac{1}{1 + q \ln(\rho)}, \quad (3.2)$$

with  $q$  and  $\rho \in (0, 1)$ , and the additional restriction that  $\rho > q$ . The latter restriction is needed to preserve the minimum of the true power law distribution on which the maximum likelihood estimator (MLE) is anchored. It is easily verified that  $\hat{\alpha}^{du}$  is always upward biased compared to Zipf's law for these parameter restrictions, implying that true inequality is underestimated. The effect of varying the parameters is also quite intuitive: an increase in  $q$  for a given  $\rho$  and a decrease in  $\rho$  for a given  $q$  increases the bias, as in

---

<sup>3</sup>Like the assumption of  $t \rightarrow \infty$  for the drift-diffusion process, the assumption of  $N \rightarrow \infty$  is necessary to make the problem analytically tractable. Our qualitative results are not materially sensitive to this assumption, as we verified by Monte Carlo simulations that the bias did not significantly differ from the large  $N$  limit in finite samples.

both cases relatively less wealth is reported for the richest.<sup>4</sup> Furthermore  $\alpha$  is only unity when either  $q = 0$  or  $\rho = 1$ , so there is no differential behavior to begin with. Thus, when it comes to underreporting, the differential behavior of the very richest compared to the relatively less wealthy is necessary to cause upward biases from the *theory first* perspective.

While we cannot derive analytical expressions for  $\rho < q$  in general, this is possible for the limit case of  $\rho = 0$ . In our stylized scenario, this would correspond to a case where the upper  $q$ -quantile is non-respondent and the wealth distribution is therefore  $q$ -truncated. In this case of *differential undersampling* or *non-response*, the richest quantile is not included at all in the sample, corresponding to a reporting rate of zero. This scenario actually appears to be empirically relevant, and connected to sampling and social desirability biases (Kennickell and Woodburn 1999; Eckerstorfer et al. 2016; Vermeulen 2016, 2018). As we show in Appendix 3.B, non-response leads asymptotically to a (strong) upward bias in the MLE of the tail exponent, now denoted by

$$\hat{\alpha}^{nr}(q) = \frac{1 - q}{1 - q + q \cdot \ln(q)}, \quad (3.3)$$

for large  $N$  and  $q \in (0, 1)$ . For this parameter range of  $q$ ,  $\hat{\alpha}^{nr}$  is always upward biased compared to the Zipf benchmark, and monotonically increasing in the quantile  $q$  of non-respondents. The quantile  $q$  of upper non- or underreporting individuals is thus the only formal difference between the competing narratives of *data first* and *theory first*.<sup>5</sup> If  $\hat{\alpha} \neq 1$ , *data first* implicitly assumes  $q = 0$  and therefore attributes all the observed deviation from the Zipf benchmark to scale-dependence in either mean or expected returns. In contrast, *theory first* takes  $\hat{\alpha} \neq 1$  to imply  $q \neq 0$  and therefore differential reporting behavior according to sample inclusion rates and the level of wealth.

### 3.3 Data and Estimation

To test the hypothesis of scale-dependence, we examine two samples covering distinct scales in the upper tail of the German wealth distribution. We need two non-overlapping samples that both exhibit power law-like top tails, as is often the case for surveys and rich lists. The German data described below comfortably meet this requirement as the minimum wealth level in the rich lists is about three times as large as the maximum wealth level in the surveys. Non-overlapping samples are necessary to isolate potential scale effects in the accumulation of wealth, and to ensure that we consider two distinct sets of wealth portfolios. The latter condition minimizes potential Type II error in hypothesis

---

<sup>4</sup>By the same token, for  $\rho > 1$ , an increase in  $q$  increases the downward bias of the estimate and thus overestimates inequality relative to the true (Zipfian) distribution.

<sup>5</sup>We would like to believe that it is not entirely trivial to reduce the impact of the two pre-analytical visions to a single parameter.

testing, because failure to reject the null hypothesis of insignificant scale differences could otherwise arise from the simultaneous presence of identical wealth portfolios, thereby affecting the estimated parameters in both sample types.

### 3.3.1 Data

The *Socio-Economic Panel* (SOEP), compiled by *Deutsches Institut für Wirtschaftsforschung*, is probably the most prominent source for microdata on German households and individuals. The 2002, 2007 and 2012 waves of the panel include items on personal wealth that we use in our analysis. Assuming different weighing and imputation techniques for the market value and disaggregation to individual values, the *SOEP* sample claims to be representative of the entire German population, implying that each person or household in Germany is chosen with equal probability (Frick et al. 2007). With a total population of 82.5 million in Germany and about 25,000 individuals in the sample, the sampling ratio thus corresponds to about 0.035 percent (Statistisches Bundesamt 2017). While the *SOEP* sample probably provides a reasonable approximation to the distribution of wealth for the majority of Germans, it is well known that wealth data from household surveys become increasingly inaccurate for the tails of the distribution (see, e.g., Davies and Shorrocks 2000). Casual empiricism indeed suggests that the reported maximum wealth level in the *SOEP* of around seventy million euros is far from being “representative” of the richest Germans, whose fortunes are about three orders of magnitude larger according to the rich lists compiled by *manager magazin*. These named lists rank the five hundred richest Germans according to their net wealth in the years 2010 to 2016. Since the rich lists are not curated for statistical inference, the data likely suffer from numerous issues regarding their consistency both in the time-series and cross-sectional domain. We discuss both datasets and their respective limitations in

### 3.3.2 Estimation

Our empirical analysis starts with the parameter estimates of the power law distributions in the upper tail of the *SOEP* and *manager magazin* samples. We interpret these as the stationary distributions resulting from a general random growth process, as described in Appendix 3.A. The assumption that the empirically observed state coincides with the stationary state of the distribution for time  $t \rightarrow \infty$  is frequently challenged though. Especially Gabaix et al. (2016) and Luttmer (2011, 2018) show that the convergence to a new stationary distribution from a shock resulting in deviations from the steady-state is extremely slow. A back-of-the-envelope calculation in Luttmer (2018) suggests that for a firm size distribution close to Zipf, but with slightly thinner tails, a shock to the aggregate capital stock would be extremely persistent with a half-life of around seventy years,



implying unrealistically low rates of recovery.<sup>6</sup> Given slow convergence, it is questionable whether the empirical distribution truly reflects the dynamics of an underlying random growth process or whether it is merely in a transient state to stationarity. On the other hand, as Levy and Levy (2003) show, the convergence to the *approximate* power law is much faster even though convergence to the *asymptotic* distribution is indeed very slow for these types of random growth processes. Levy and Levy (2003) understand approximate convergence as convergence to a distribution that cannot be statistically distinguished from the stationary state by means of a Kolmogorov-Smirnov (KS) test. If the parameter estimates are at least approximating the true stationary state of the random growth process, the pronounced differences we find between samples will not be mere artefacts of one distribution being in a transient state but not the other, thus reflecting genuine differences in reporting, sampling, or the underlying growth process.

Given the diffusion in Appendix 3.A, we reject the null hypothesis of scale-independence for  $\hat{\alpha}$  significantly different from unity. This procedure is advantageous in the sense that it relies on observables to test for scale-dependence and thus allows inferences about the (at least partially) unobservable random growth process. Additionally, we also consider the distribution of growth rates in wealth to judge whether scale-(in)dependence characterizes the wealth accumulation process. Notice that the diffusion in Appendix 3.A requires us to consider scale-dependence in both expected value and risk, which are readily measured by the MLEs of the location and dispersion parameters of the growth rate distributions. We estimate the tail exponent of the power law using maximum likelihood.<sup>7</sup> For the estimation of the minimum wealth level  $w_{min}$  in the *SOEP* sample, we use the standard suggested by Clauset et al. (2009), yielding  $\hat{w}_{min} = 280,000$  euros for the 2002 sample,  $\hat{w}_{min} = 200,000$  euro for the 2007 sample, and  $\hat{w}_{min} = 180,000$  euro for the 2012 sample.<sup>8</sup> It seems reasonable to assume that a net worth of around 200,000 euros already gives rise to primarily multiplicative returns, especially considering that most households hold their wealth in the form of owner-occupied housing. Since the rich lists should be characterized by power laws, we do not estimate  $w_{min}$  but rather take it directly from the data, so  $w_{min}$  simply corresponds to the minimal wealth level in each rich list, ranging from 200 to 250

---

<sup>6</sup>For a distribution that is exactly Zipf, there would be no recovery at all.

<sup>7</sup>Clauset et al. (2009) show that an MLE fit is the least biased method to estimate the tail index (or *characteristic exponent*) of a power law, compared to OLS methods or a linear fit on double-logarithmic scale. See Goldstein et al. (2004) for a more rigorous analysis of different graphical methods and their respective shortcomings. Even though we estimate from a discrete dataset, we estimate the power law for its continuous analogue, as the analytical results for different biases in the results section are based on the continuous version. The continuous MLE is introduced and discussed in Appendix 3.B in more detail. Our estimates are not materially sensitive to the choice between the discrete and continuous estimator. Standard errors are determined by exploiting the asymptotic normality of the MLE (De Haan and Resnick 1997).

<sup>8</sup>Clauset et al. (2009) show that their method, based on iteratively increasing the reverse order statistic until a goodness-of-fit test (like the KS test) rejects a power law distribution, outperforms other possible procedures, such as minimizing the Bayesian information criterion (BIC).

million euros in the different years.<sup>9</sup> The minimum in the rich lists is thus three orders of magnitude larger than in the surveys.

Finally, we would expect the distribution of wealth growth rates to be Laplacian (or double-exponential) since we measure wealth growth by the logarithmic difference in wealth levels, that is  $r_{i,t} = \log(w_{i,t}) - \log(w_{i,t-1})$  for agent  $i$  during the period  $t$  to  $t - 1$ . It can be shown that  $\log(w)$  follows an exponential distribution if  $w$  follows a power law, and that the difference between two exponentially distributed variables is Laplacian (Kotz et al. 2001). The symmetric Laplace distribution for returns  $r$  then has a probability density function (PDF) that is given by

$$f(r; m, \sigma) = \frac{1}{2\sigma} e^{-|\frac{r-m}{\sigma}|}, \quad (3.4)$$

where  $m \in \mathbb{R}$  and  $\sigma > 0$  are location and dispersion parameters, respectively.<sup>10</sup> From a conventional point of view,  $m$  measures the expected return in a set of wealth portfolios, while  $\sigma$  measures the associated risk in these portfolios.

Our estimation strategy considers the cross-sectional distributions of wealth in both samples, each interpreted as the outcome of a parsimonious random growth process like the one described in Appendix 3.A, whose realizations are at least partially unobservable. The estimated tail index then allows us to infer scale-(in)dependence within this unobservable process. Moreover, using the Laplace estimates from the actual growth rate distributions, we can test parametrically for scale-(in)dependence in expected returns or risk, and we also employ several nonparametric tests.

### 3.4 Results

Our parametric estimation strategy is based on the two distributional regularities in the upper tail of cross-sectional wealth portfolios, namely the power law distribution in wealth levels and the Laplace distribution of portfolio returns, because the respective empirical densities are reasonably in line with the theoretically expected functional forms. The observed complementary cumulative distribution functions (complementary CDFs) above the minimum thresholds  $w_{min}$  are approximately linear on a double-logarithmic scale, indicating power law-like patterns for the richest individuals in both samples (see The empirical densities of returns to wealth portfolios are also reasonably well approximated by the expected Laplace distribution. This is readily indicated by their (symmetric)

---

<sup>9</sup>This procedure is also advantageous in so far as the respective  $w_{min}$  levels ensure that the estimated power laws always span at least two orders of magnitude, usually considered to be a minimum requirement for significantly claiming a power law distribution in the first place (Stumpf and Porter 2012).

<sup>10</sup>In a strict sense, symmetry of the Laplace distribution is not guaranteed because the parameter values of the power law tail might be time-varying. The correct distribution would then be an asymmetric Laplace distribution of wealth returns, yet we find the empirical distributions to have skewness that is not statistically different from zero, indicating that the symmetric version is empirically useful.

tent shape on semi-logarithmic scale that is characteristic of the Laplace, and shown in Appendix 3.C . Apart from mere visual inspection, the standard procedure to test for a Laplace distribution is to fit an exponential power (or Subbotin) distribution to the data (Subbotin 1923). Since the Subbotin distribution includes the Laplace as a special case when its shape parameter equals unity, an MLE fit of the Subbotin parameters provides a convenient test. As we show in Appendix 3.C , a shape parameter of unity cannot be rejected in any of the considered cases, so our findings should not be distorted by systematic deviations from the parametric forms we impose for the estimations.

### 3.4.1 Distributional Results

We estimate the parameters for the power law distribution separately for the survey and the rich list. The tail indices are estimated via maximum likelihood, employing the respective empirical minima from the rich list, and using the procedure described in Clauset et al. (2009) to estimate the respective minima  $\hat{w}_{min}$  in the survey. Tables 3.1 and 3.2 report tail index estimates for the survey tails and rich lists, respectively.

<i>SOEP</i>	2002	2007	2012
Tail exponent estimate $\hat{\alpha}$	1.3144 (0.0423)	1.0978 (0.0324)	1.2982 (0.0354)
Minimum wealth level estimate $\hat{w}_{min}$	0.28	0.20	0.18
Maximum wealth level $w_{max}$	70.55	30.60	16.00
Sample size $N$	961	1,260	1,332

Table 3.1: Tail index estimates  $\hat{\alpha}$  for the power law region of the *SOEP* survey with standard errors in parentheses. The minimum and maximum wealth levels  $\hat{w}_{min}$  and  $w_{max}$  are stated in millions of euros, deflated with index year 2010, while  $N$  denotes the number of observations in the power law tail.

Two peculiarities stand out. First, normality of standard errors for the tail index estimates (De Haan and Resnick 1997) implies that Zipf’s law (with  $\alpha = 1$ ) can be rejected with at least 95% confidence in all survey years. Wealth in the survey tails therefore appears more equally distributed than scale-independent growth would imply. Second, the wealth maxima in the survey are not even on the same order of magnitude as the wealth minima reported in the rich lists. These implausibly small maxima indicate severe undersampling (or rather the complete absence) of the super-rich in the survey, and are a major reason for the relatively low degree of measured inequality in the survey tails. Tail index estimates for the rich list, on the other hand, stand in stark contrast to those for the survey. As shown in Table 3.2, we cannot reject Zipf’s law at the usual significance levels in any of the years other than 2013 and 2016, with Zipf’s law being only barely

<i>manager magazin</i>	2010	2011	2013	2014	2015	2016
Tail exponent estimate $\hat{\alpha}$	0.9983 (0.0450)	0.9863 (0.0442)	1.0999 (0.0495)	0.9874 (0.0443)	0.9358 (0.0419)	0.7615 (0.0341)
Minimum wealth level estimate $\hat{w}_{min}$	200	200	250	250	250	200
Maximum wealth level $w_{max}$	17,100	19,000	23,950	31,000	26,500	30,000
Sample size $N$	499	498	494	497	500	500

Table 3.2: Tail index estimates  $\hat{\alpha}$  for the *manager magazin* rich list with standard errors in parentheses. The minimum and maximum wealth levels  $w_{min}$  and  $w_{max}$  are stated in millions of euros, deflated with index year 2010, while  $N$  denotes the number of observations in the power law tail.

rejected in 2013.<sup>11</sup> In the language of the stochastic accumulation process, Zipf’s law indicates scale-independent returns among the super-rich. Yet significant deviations from Zipf’s law in the survey tails point to scale-dependent wealth returns within the survey populations, and obviously also to scale-dependent returns between the two sample types. Hence we consider the distribution of wealth returns in the two sample types, and to facilitate comparison we construct wealth returns over five year intervals. Several non-parametric tests reject the null hypothesis of distributional equivalence between the two sample types in both periods, but fail to reject it within the samples between periods (cf. Apparently the data suggest that wealth dynamics are time-invariant but scale-dependent between sample types. The parameter estimates for the Laplace distribution of wealth returns, summarized in Table 3.3, strengthen the impression from the non-parametric tests. The estimates for the location parameter  $m$  (the “average” return) and the dispersion  $\sigma$  (the “average” risk) do not vary much within the respective samples, yet vastly differ between the two sample types. While average returns do not significantly differ from zero in the survey tails, zero can safely be rejected at the five percent level in the rich lists, where  $m$  is significantly positive, implying that Germany’s super-rich on average became wealthier during the considered period. Paradoxically, however,  $\sigma$  is significantly lower in the rich lists than in the survey tails, apparently indicating that super-rich portfolios are less risky than the ones in the survey tails.<sup>12</sup> So how can we interpret these findings?

### 3.4.2 Data First

Taken at face value our findings indicate that the accumulation process is scale-dependent in the survey study but scale-independent in the rich list, where we find higher average returns and lower volatility compared to the survey. From a theoretical point of view this

<sup>11</sup>According to *manager magazin* staff, 2016 is the only year in which they tried to account for wealth held in foundations or charitable organizations. Thus the 2016 tail index does not measure the same concept, which is why we mostly discard 2016 in our analysis.

<sup>12</sup>This interpretation, although entirely conventional, needs to assume ergodicity in returns to wealth, which is why we use quotation marks for the notions of “average” return or “average” risk.

Laplace parameter estimates	$\hat{m}$	$\hat{\sigma}$
<i>manager magazin</i> 2010–15	0.1024 (0.0090)	0.1458 (0.0010)
<i>manager magazin</i> 2011–16	0.0824 (0.0117)	0.1319 (0.0082)
<i>SOEP</i> 2002–07	0.0280 (0.0144)	0.3710 (0.0077)
<i>SOEP</i> 2007–12	0.0215 (0.0149)	0.3745 (0.0076)

Table 3.3: Maximum likelihood parameter estimates for the Laplace distribution of wealth returns, with standard errors in parentheses. While a location measure or “average” return of zero cannot be rejected for the survey tails, it is significantly greater than zero in the rich lists. Note that the dispersion of returns, that is the “average” risk across portfolios, is markedly lower in the rich lists than in the survey tails.

is puzzling. How plausible is it that the investment strategies of the super-rich converge to roughly equivalent risk profiles that not only outperform other (still rather) wealthy individuals, but do so at a lower risk? The conventional rationale for the risk-return tradeoff, as for instance in the canonical intertemporal capital asset pricing model of Merton (1973), suggests that the conditional expected excess return should grow linearly with its conditional variance.<sup>13</sup> But both the non-parametric tests as well as the parameter estimates for the Laplace distribution of wealth returns indicate that the super-rich enjoy higher expected returns at lower risk. The excess returns of Germany’s super-rich cannot be explained by a higher risk tolerance, because this should be reflected in a higher dispersion of returns among the super-rich.

The *data first* interpretation thus not only suggests scale-dependence, but scale-dependence that cannot be explained by heterogeneous risk preferences alone. To explain the estimation results within the framework of random multiplicative growth, we need to assume that financial markets are not fully competitive in the conventional sense. This would suggest that investors’ talent or the increased set of possibilities that comes with being very wealthy enables the richest to persistently beat the market and achieve above average risk-adjusted returns at a lower risk. Such an interpretation would also be at odds with empirical findings on risk preferences that observe decreasing risk-aversion in wealth levels, such that higher net worth correlates positively with a higher dispersion in returns to wealth (Guiso et al. 1996; King and Leape 1998; Calvet and Sodini 2014). The *data first* interpretation thus poses a challenge to both, the empirically observed risk profiles,

<sup>13</sup>There exists an ongoing debate on whether this relationship can be established empirically (French et al. 1987; Campbell 1987; Nelson 1991; Campbell and Hentschel 1992; Harvey 2001; Goyal and Santa-Clara 2003; Brandt and Kang 2004; Ghysels et al. 2005; Bali and Peng 2006; Andersen et al. 2006; Guo and Whitelaw 2006; Lundblad 2007; Bali 2008; Gonzales et al. 2012). While most of the studies find at least weak support for the risk-return trade-off for various time frames and markets, the debate seems to have now shifted to the precise functional form of this relationship - as opposed to the linear one implied by the CAPM.

and the idea that financial markets with rapid feedbacks and a low degree of informational asymmetries should be close to the benchmark of a fully competitive market.

### 3.4.3 Theory First

Our central point here is that these “puzzles” can be resolved within the *theory first* interpretation once we agree that estimates of the tail exponent in the two samples suffer from two different sources of bias. Equiprobable sampling in the survey makes it very unlikely to observe the largest wealth levels that are necessary for reliable estimation of the tail index, as we quantify in Appendix 3.E. Note that the probability of including the maximum wealth level for the *SOEP* sampling ratio under equiprobable sampling is 0.035 percent and thus practically equal to zero. Adding to this problem are concerns of social desirability biases, particularly the phenomenon that the super-rich tend not to respond to survey requests. As the probability of non-response is therefore positively correlated with wealth levels, the survey is subject to *differential non-response* (Kennickell and Woodburn 1999; Eckerstorfer et al. 2016; Vermeulen 2016, 2018). These two considerations lead to *undersampling*, that is the largest wealth levels are not included at all in the survey sample. In contrast, the rich list is a carefully selected sample aimed at covering the super-rich, and one can therefore expect that undersampling is not an issue. On the other hand, the *manager magazin* staff relies on public records for their compilation of the rich list, likely underestimating the actual wealth levels for Germany’s super-rich due to privacy considerations and tax avoidance that is particularly pronounced among the wealthiest (Alstadsæter et al. 2019). Consequently, the *manager magazin* sample should be subject to *underreporting*, not undersampling. In more colloquial terms, the upward bias in the survey sample arises because the richest are not included at all in the sample, while the upward bias in the rich list arises because the richest are not included with the full extent of their wealth.

To study the relative biases arising from differential undersampling and underreporting, we plot the upward deviation from the theoretically expected tail exponent of  $\alpha = 1$  for different reporting rates  $\rho$  and undersampling in the (empirically motivated) quantile  $q \in (0, 0.2)$ . The case  $\rho = 0$  corresponds to undersampling, and is also the only case for  $\rho < q$  that we can examine along the lines elaborated in section 3.2.

Figure 3.1 supports the intuition that the relative bias is decreasing in the reporting rate  $\rho$ , since for smaller  $\rho$  a larger fraction of wealth is not reported. When  $\rho = 1$  we recover the initial distribution from eq. (3.2), and there is no bias for any  $q$ . Compared to the underreporting bias, the undersampling bias is rather unexpected though. If merely 25% of wealth were to be reported by the richest  $q$ -quantile, this would lead to a disproportionately smaller bias in the estimator, indicating that tail index estimates from the rich list are in all likelihood much less (upward) biased than estimates from the survey. This

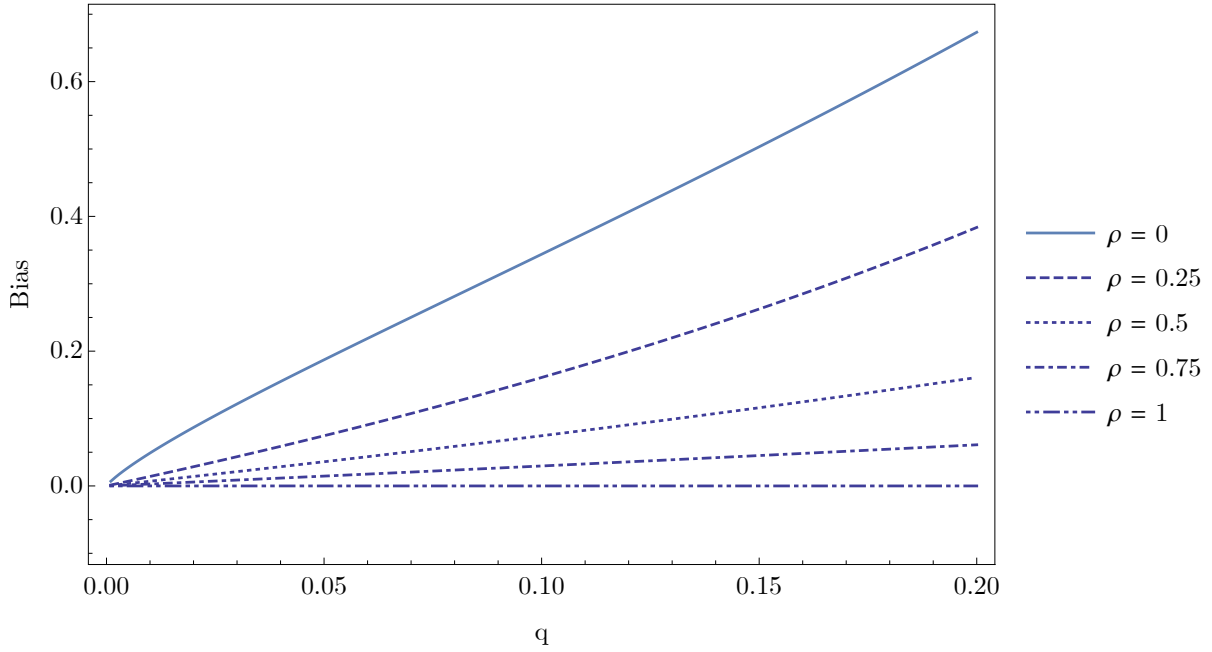


Figure 3.1: Illustration of the analytical bias resulting from different combinations of reporting rates  $\rho$  and underreporting fractions  $q$  compared to the theoretically expected  $\alpha$  of unity for Zipf’s law. The strength of the bias increases disproportionately with decreasing reporting rates, where  $\rho = 0$  leads to much more upward biased tail index estimates, even compared to  $\rho = 0.25$ .

is reminiscent of the finding by Cristelli et al. (2012) that the maximum in a power law is most informative. Our result is more general in the sense that even partial inclusion of these top observations by only a fraction of their true level will greatly reduce the bias in measured inequality. Given the limited impact of differential underreporting, we conclude that the true inequality of the system is substantially closer to the Zipf benchmark than the survey estimates seem to suggest, as indicated by the less biased estimates for the rich list.

Furthermore, our closed-form expression (3.2) that quantifies the impact of underreporting on the tail exponent also allows us to back out the reporting rates  $\rho$  for the rich lists. We assume that the upper 20% quantile exhibits different reporting behavior, in the sense that the richest one hundred Germans constitute a rather salient set on the rich list, where the *manager magazin* staff focuses their efforts to compile reliable data (Balz et al. 2014), and thus the effect of tax avoidance should not be compounded by rounding errors or selection bias for the considered sources.<sup>14</sup> So fixing  $q = 0.2$  and further assuming that Zipf’s law governs the true distribution, we obtain the reporting rates  $\rho$  in Table 3.4.

The implied reporting rates appear to be plausible, except for the 2016 estimate that neatly reflects the qualitative change in the data collection procedure by the *manager*

<sup>14</sup>This is supported by apparent “digit-preferences” or “heaping effects” that we can observe below rank 100, where the data seem abnormally clustered in increments of fifty million euro (Heitjan and Rubin 1991; Schneeweiß et al. 2010).

<i>manager magazin</i>	2010	2011	2013	2014	2015	2016
Implied reporting rate $\rho$	1.0086 (0.1959)	1.0719 (0.2095)	0.6350 (0.1517)	1.0659 (0.2083)	1.4092 (0.2884)	4.7878 (1.1743)

Table 3.4: Implied reporting rates  $\rho$  for differential underreporting in the rich lists, with standard errors in parentheses. For illustrative purposes, we assume that  $q = 0.2$  and that the true distribution follows Zipf’s law exactly. Except for 2016, the estimates suggest little differential reporting behavior, adding to the plausibility of the *theory first* interpretation.

*magazin* staff.<sup>15</sup> Note the highly non-linear and perhaps counterintuitive effect of a mere twenty percent decrease in  $\hat{\alpha}$  between 2015 and 2016 that requires the implied reporting rate to more than triple, showing that the change of sampling procedures between 2015 and 2016 is a qualitative shift that would easily be missed if we were to exclusively look at the twenty percent increase in measured inequality. Thus the assumption of Zipf’s and consequently Gibrat’s law along with scale-independence seem entirely plausible in the *theory first* interpretation, especially since we cannot reject Zipf’s law in any of the years other than 2013 and 2016.

Regarding the survey, *theory first* suggests that the deviation from Zipf’s law in the *SOEP* data originates from undersampling ( $\rho = 0$ ) such that the super-rich are entirely absent in the sample. Using the closed-form expression (3.3), we can infer the  $q$ -quantiles of non-respondents from the estimated tail exponents both in the survey tail, denoted  $q_{pl}$ , and also for the survey as a whole,  $q_{tot} = q_{pl}(n_{pl}/n_{tot})$ , where  $n_{tot}$  denotes the size of the *SOEP* sample and  $n_{pl}$  denotes the size of the survey tail (reported in the last row of Table 3.1). The results are summarized in Table 3.5.

<i>SOEP</i>	2002	2007	2012
Implied non-response rate PL tail $\hat{q}_{pl}$	0.0906 (0.0135)	0.02310 (0.0094)	0.08538 (0.0114)
Implied non-response rate full sample $\hat{q}_{tot}$	0.0030 (0.0005)	0.0010 (0.0004)	0.0039 (0.0005)

Table 3.5: Implied non-response rates in the survey tail,  $\hat{q}_{pl}$ , and in the entire survey,  $\hat{q}_{tot}$ , calculated from eq. (3.3) under the assumption of Zipf’s law, with standard errors in parentheses. Non-response rates are tiny and imply that missing merely twenty-five to a hundred of the super-rich in the survey can already explain the observed deviations from Zipf’s law.

The implied non-response rates relative to the size of the survey sample are remarkably low. The effects of equiprobable sampling combined with differential non-response quite

<sup>15</sup>Reporting rates  $\rho > 1$  could, at least in principle, arise from the salience of the richest quantile through intensified compilation efforts that lead to a relative overestimation of top wealth. So whenever  $\rho > 1$ , the salience bias would outweigh the tax avoidance and social desirability biases. More importantly, however, a reporting rate of  $\rho = 1$  can only be safely rejected in 2016, while in 2013 we are just very slightly above the conventional five percent level. In all other years, we cannot reject  $\rho = 1$  at this significance level.



plausibly lead to non-response rates  $q_{tot}$  of 0.1 to 0.4%. Consequently, the survey data are not inconsistent with the interpretation of scale-independent multiplicative growth and therefore Zipf's law in wealth levels. Since the mixture of non-overlapping Zipf samples is also distributed as a Zipf law, the *theory first* interpretation supports scale-independence across the entire tail of the German wealth distribution. After all, our results underline the importance of maximum wealth levels for the estimation of tail indices because failing to account for merely 0.1 to 0.4 % of the richest individuals already leads to substantial biases—and the descriptive statistics for the two samples clearly indicate that the actual response rate of the super-rich in the survey is zero. As we show in the upcoming subsection, differences in tail index estimates translate into substantial differences in estimated top tail wealth, and therefore also lead to enormous differences in measures of wealth inequality.<sup>16</sup>

### 3.4.4 Total Wealth Estimates

How much wealth is concentrated in the power law tail? The most recent literature on this matter extrapolates the estimates from survey studies to a maximum determined from rich lists (Vermeulen 2018; Bach et al. 2019). Even within this established methodology, three very different kinds of answers emerge depending on the pre-analytical vision one employs. In line with the literature, we use the continuous analogue of the power law distribution and integrate to derive a measure for the total power law wealth  $W$ . The minima correspond to the estimates for the survey study, while we take the maxima from the rich lists. Within the *data first* interpretation, we need to choose between the estimated tail exponents from the survey study and the rich list corresponding to the respective belief that either the inequality within the *SOEP* or the *manager magazin* sample is more representative of the power law tail as a whole. The *theory first* perspective suggests Zipf's law and thus leaves no such degree of freedom. The estimation strategy is elaborated in more detail in Appendix 3.D, where we also detail how to estimate the population  $n$  inhabiting the power law tail.

In the *data first* estimations, we essentially extrapolate the power law population in-sample to the entire German population of  $N = 82.5$  million (Statistisches Bundesamt 2017). This simple extrapolation is justified since *data first* assumes no systematic non-response rates for the richest. The estimates for the population from the survey study reveal a relatively large power law population with a relatively homogeneous wealth distribution, while the estimates for the rich list imply a very small population characterized by an extremely heterogeneous wealth distribution. The *theory first* perspective implies Zipf's law for the entire top tail and attributes observed differences from this benchmark in the survey to differential non-response. We thus correct our population estimates for

---

<sup>16</sup>This will of course also be true for measured inequality with respect to the entire population, and not just within the power law tail that we focus on here.

the survey by the estimated non-response rates. Unanimously, we find the largest estimated power law populations for this *theory first* perspective (see Appendix 3.D ). Both the corrected as well as the uncorrected estimates for the survey study differ up to one order of magnitude with respect to the estimates for the rich list. The correction within the *theory first* approach has a very limited effect on the estimated total population, resulting from the fact that the estimated non-response rates are tiny. This leaves us with three estimation strategies, each with 18 possible combinations of  $\hat{w}_{min}^{SOEP}$  and  $w_{max}^{mm}$  for all sampling years. Table 3.6 shows how the differences in the estimated power law populations and tail indices translate into differences in total wealth.

<hr/> <hr/>						
mm $\hat{W}$						
SOEP years / mm years	2010	2011	2013	2014	2015	2016
2002	1,096	1,079	1,819	1,409	1,280	1,171
2007	1,129	1,109	1,910	1,447	1,305	1,177
2012	1,140	1,118	1,939	1,458	1,313	1,179
SOEP $\hat{W}$						
SOEP years / mm years	2010	2011	2013	2014	2015	2016
2002	3,114	3,117	3,124	3,131	3,126	3,130
2007	5,420	5,447	5,506	5,571	5,532	5,563
2012	2,884	2,887	2,894	2,900	2,896	2,900
Zipf $\hat{W}$						
SOEP years / mm years	2010	2011	2013	2014	2015	2016
2002	8,498	8,579	8,757	8,956	8,835	8,931
2007	8,185	8,261	8,428	8,614	8,500	8,590
2012	7,882	7,955	8,114	8,292	8,184	8,269
<hr/>						

Table 3.6: Estimated wealth in the power law tail for combinations of minima and maxima from the respective survey and rich list samples in billions of euros (inflation-adjusted with base year 2010). Details regarding the underlying estimation strategies and parameter constellations are described in Appendix 3.D . The estimates exhibit tremendous variation, almost spanning one order of magnitude.

We note first that especially the 2007 estimates for the *SOEP* are in remarkably close agreement with the latest estimates in Bach et al. (2019) based on the Household Finance and Consumption Survey (HFCS), even though our samples differ substantially from theirs.<sup>17</sup> This is also the case where the estimation procedure for the total power law population is closest to theirs. We take this as evidence that our results are not driven by idiosyncrasies in our data and instead testify to the external validity of our approach. Second, and more importantly, the results differ enormously between the two pre-analytical visions. The estimates for the pure Zipf case are higher than the rich list estimates by at least a factor of six, in some cases even by one order of magnitude. This is primarily caused by the huge differences in estimated population, with both estimation strategies appearing to be plausible. Even when populations are not differing too much,

<sup>17</sup>The main difference is that the estimation by Bach et al. (2019) can exploit survey weights of the HFCS for oversampling which are not available for the *SOEP*.

the pre-analytical vision has a large effect on estimated total wealth, as the uncorrected and corrected estimates for the pure survey and Zipf case show, differing by up to a factor of three. Hence even state-of-the-art methods for this kind of estimation will likely severely underestimate the degree of inequality both within the richest group, and also between the top tail and the rest of the population. A case can be made (more or less convincingly) for all three estimation strategies, and it seems fair to say that total wealth estimates are influenced at least as much by pre-analytical belief as they are by the data used for estimation.

## 3.5 Discussion

We have shown that the pre-analytical vision decisively informs the research agenda as well as the conclusions drawn from it. So how wealthy are the rich, and are returns to wealth scale-dependent for them? As we have argued here, the proposed mutually contradicting interpretations of *data first* versus *theory first* are observationally equivalent to each other. *Data first* interprets the observed deviations from Gibrat's law in wealth returns, and consequently Zipf's law in wealth levels, as evidence for scale-dependence. *Theory first*, on the other hand, explains these deviations through sampling and reporting biases that affect the two sample types differently. Ultimately, we cannot discriminate between the two narratives based on the data alone, and seem to face a classic instance of the underdetermination of scientific theory by evidence, featuring prominently in the philosophy of science at least since the turn of the 20th century (Quine 1975). On the other hand, dearly held convictions in economic theory, such as the risk-return trade-off, informationally efficient markets, and the classical notion of competition that requires an equalization of rates of return, patently suggest that *theory first* is a more plausible explanation for the data.

The proposed differential biases cast doubt on the validity of conclusions drawn across and within sample types, both in the cross-sectional and the time series domain. Valid inference in the presence of reporting biases requires stability of parameters over time, otherwise identified trends might become spurious and instead reflect changes in bias. Since the proposed explanation of biases is behavioral and builds on empirically well-established phenomena such as salience, differential tax avoidance, or social desirability rather than being based on sampling method, there is no reason to expect stability. The estimated parameters within the *theory first* framework indeed suggest such variable behavioral responses over time.

Improving data availability and quality, for instance through the use of wealth or capitalized income tax data, might mitigate the severity of undersampling. Data availability then depends on the political willingness to impose such taxes in the first place, while future research will still be confined to the taxed population and a legal definition of

wealth that is generally not catered towards the needs of statistical inference (Galbraith 2019). Our results thus highlight the need to improve on survey and sampling methods, not to abandon them altogether. The recently conducted *SOEP-P* sample, which uses information on stock holdings to target high net worth individuals, is a first step in this direction, but still fails to adequately capture the super-rich in the targeted random sample, and thus fails to include the maximum wealth levels that we show to be crucial for valid inference. This is why data from the *SOEP-P* need to be complemented by the *manager magazin* rich lists in the hope of adequately capturing tail wealth. Yet our findings strongly suggest that the *SOEP-P* supplement and the resulting composite sample from the rich lists still must be scrutinized along the lines of the fundamental *theory first* versus *data first* distinction. In the end, our results cast serious doubt on simply pooling data from different sample types and comparing trends therein, which has been standard practice so far (see, e.g. Vermeulen 2018; Bach et al. 2019; Schröder et al. 2020).

While it is not surprising that the two narratives yield different estimates for total wealth in the top tail, the magnitude of this difference comes probably unexpected for most, because power laws have the counter-intuitive property that supposedly small variations in the tail index lead to enormous variations in totals. This property is substantially compounded by tiny degrees of undersampling, here on the order of a tenth of a percent, that lead to differences in estimated total wealth by a factor of up to three. Such small degrees of undersampling are easily explained by equiprobable sampling from a power law, leading to an inclusion probability of the maximum that is on the order of a hundredth of a percent in our case, and thus practically equal to zero. Since we have shown how important the inclusion of an accurately measured maximum is for unbiased tail index estimation, this is disconcerting.

The enormous differences between total wealth estimates suggest that inferences from survey studies regarding the cross-sectional distribution of wealth and its time variation tend to be severely distorted, illustrating that discussions about the notion of “representativeness” in scale-free systems are not discussions about technical subtleties but disagreements in substance. In Appendix 3.F, we conduct a simple analytical thought experiment for an extreme case of unrepresentative oversampling of the rich using *logarithmic sampling* across different orders of magnitude in wealth levels. We show that for Zipf’s law, the necessary sampling ratio to surely include the maximum decays extremely fast by a power function. If the maximum is indeed as important as our analytical results on the undersampling bias indicate, even conventional oversampling techniques will be insufficient, and should instead try to implement logarithmic sampling in order to allow for unbiased estimations. After all, our results show that accurate representations of total wealth require us to be highly “unrepresentative” in the sampling of individuals.

### 3.A Model

Consider the Markov diffusion with support over the real half-line  $(0, \infty)$  for the normalized wealth  $w$  of a typical household or individual given by

$$dw_t/w_t = \mu(w)dt + \sigma(w)dW_t, \quad (3.5)$$

where  $\mu$  is the mean growth rate of normalized wealth,  $\sigma$  its standard deviation and  $dW_t$  are Wiener increments. Denote by  $f(w, t)$  the distribution of normalized wealth levels at  $t$ , and by  $f(w)$  the stationary density for  $t \rightarrow \infty$ . The Fokker-Planck equation is then given by

$$\frac{\partial}{\partial t} f(w, t) = -\frac{\partial[\mu(w)wf(w, t)]}{\partial w} + \frac{1}{2} \frac{\partial^2[\sigma^2(w)w^2 f(w, t)]}{\partial w^2}. \quad (3.6)$$

For the stationary state, it has to hold that

$$0 = -\frac{\partial[\mu(w)wf(w)]}{\partial w} + \frac{1}{2} \frac{\partial^2[\sigma^2(w)w^2 f(w)]}{\partial w^2}. \quad (3.7)$$

Integration yields

$$0 = -[\mu(w)wf(w)] + \frac{1}{2} \frac{\partial[\sigma^2(w)w^2 f(w)]}{\partial w}. \quad (3.8)$$

This allows to solve for  $f(w)$  by differentiating the right term and omitting the dependence on  $t$  for the stationary density by

$$0 = -\mu(w) \cdot w \cdot f(w) + \frac{1}{2} \left[ \frac{\partial \sigma^2(w)}{\partial w} \cdot w^2 \cdot f(w) + \sigma^2(w) \cdot 2w \cdot f(w) + \sigma^2(w) \cdot w^2 \cdot f'(w) \right] \quad (3.9)$$

and therefore

$$f(w) = \frac{\sigma^2(w) \cdot w^2 \cdot f'(w)}{2\mu(w)w - (\partial \sigma^2(w)/\partial w)w^2 - 2\sigma^2(w)w}. \quad (3.10)$$

Establishing conditions for stationarity or convergence to a power law distribution is far from trivial. Informed by his application to city sizes, Gabaix (1999) assumed both a time-invariant population size  $N$  and minimum level  $w_{min}$ , unaware of the result in Blank and Solomon (2000) that the distribution approaches a degenerate case with  $\alpha = 0$  if both variables are held constant. To guarantee convergence to a stationary power-law, we follow Malcai et al. (1999) and Blank and Solomon (2000) and assume a time-invariant population size  $N$  and a time-varying reflecting boundary  $w_{min}(t)$  that depends on the

average wealth  $\bar{w}(t)$  by some small constant  $c \in \mathbb{R}^+$ . The minimum threshold to “join the super-rich” should therefore increase over time, at the very least through inflation in the monetary value of wealth portfolios, so the latter assumption does not appear too restrictive to be of general interest. Under these assumptions, the tail exponent of the stationary density  $\alpha$  is

$$\alpha(w, f) = \frac{-w \cdot f'(w)}{f(w)} - 1. \quad (3.11)$$

Substituting equation (3.10) in (3.11) yields

$$\alpha(\mu, \sigma) = 1 - 2 \frac{\mu(w)}{\sigma^2(w)} + \frac{w}{\sigma^2(w)} \frac{\partial \sigma^2(w)}{\partial w}. \quad (3.12)$$

Since we consider normalized wealth levels,  $\mu(w)$  corresponds to the excess expected growth rate relative to the average growth rate across all wealth levels  $\bar{\gamma}$  by  $\gamma(w) - \bar{\gamma}$ , implying

$$\alpha(\mu, \sigma) = 1 - 2 \frac{\gamma(w) - \bar{\gamma}}{\sigma^2(w)} + \frac{w}{\sigma^2(w)} \frac{\partial \sigma^2(w)}{\partial w}.^\dagger \quad (3.13)$$

Zipf’s law emerges as a special case of growth rates characterized by Gibrat’s law. This implies that the partial  $\partial \sigma^2(w)/\partial w$  is zero, as there is no scale-dependence in the variance. Also, Gibrat’s law implies that the expected normalized growth rate, that is, the excess growth rate of wealth levels  $w$  in relation to the average growth rate, is independent of  $w$  for any  $w$  and thus must be zero, thereby implying the Zipf exponent of  $\alpha(0, \sigma) = 1$ . To confirm this, consider the general diffusion in equation (3.5), with  $\mu(w) = \mu = 0$  and  $\sigma(w) = \sigma$ . The general Fokker-Planck equation (3.6) under these assumptions is

$$\begin{aligned} 0 &= -\frac{\partial[0 \cdot w \cdot f(w)]}{\partial w} + \frac{1}{2} \frac{\partial^2[\sigma w^2 f(w)]}{\partial w^2} \\ &= \frac{1}{2} \frac{\partial^2[\sigma w^2 f(w)]}{\partial w^2}. \end{aligned} \quad (3.14)$$

It is easy to see that a density  $f(w)$  solves equation (3.14), whenever the differentiated term in (3.14) is independent of  $w$ . This is exactly the case for  $f(w) = C/w^2$ , that is, Zipf’s law with a normalizing constant  $C$  independent of  $w$ .

---

<sup>†</sup>In Gabaix (1999), there is a minor typographical error on page 757, where the correct expression for the tail exponent in equation (13) should read  $\gamma(S)$ , not  $\zeta(S)$ , like for our analogous expression in equation (3.13).

### 3.B Analytical Results for the Estimation of $\alpha$ under $q$ -Truncation

**Zipf's Law and Hill Estimator.** Preliminaries.

Suppose a discrete quantity  $w$  is distributed according to Zipf's law, so its tail index  $\alpha$  equals unity. According to the *rank-size* formulation, its values are therefore given by

$$w(s) = \frac{w_{\max}}{s}, \quad (3.15)$$

with  $s = 1, 2, \dots, N$  as the respective ranks of a given  $w$  in descending order,  $N$  as the number of values with  $N \in \mathbb{N}^+$ , and  $w_{\max}$  as the maximum value of the distribution. Equivalently, rewriting equation (3.15) in terms of the minimum value  $w_{\min}$  yields

$$w(s) = \frac{w_{\min} \cdot N}{s}, \quad (3.16)$$

since for Zipf it holds that  $w_{\max} = N \cdot w_{\min}$ . *Maximum likelihood estimation* (MLE) for any given (continuous) power law yields the *Hill estimator* (Clauset et al. 2009), that is

$$\hat{\alpha}(w_{\min}; N) = N \cdot \left( \sum_{s=1}^N \ln \left( \frac{w(s)}{w_{\min}} \right) \right)^{-1} \quad (3.17)$$

which is by equation (3.16)

$$= N \cdot \left( \sum_{s=1}^N \ln \left( \frac{N}{s} \right) \right)^{-1}, \quad (3.18)$$

now independent of  $w_{\min}$  and converging asymptotically for  $N \rightarrow \infty$  to Zipf's law, that is  $\alpha = 1$ .

**Unanimous Proportional Underreporting.** Unbiasedness Result.

Suppose that a discrete quantity is perfectly distributed according to Zipf's law. All individuals report only a fraction  $\rho \in (0, 1)$  of this quantity (the response-rate), which implies unanimous (proportional) underreporting. The rank-size rule for unanimous underreporting thus reads

$$w^{uu}(s) = \rho \cdot \frac{w_{\min} \cdot N}{s}, \quad (3.19)$$

with  $s = 1, 2, \dots, N$  as the ranks. Notice that we also require  $w_{\min}^{uu} = \rho \cdot w_{\min}$  by  $w(N) = w_{\min}$  and  $w^{uu}(N) = w_{\min}^{uu}$ . The Hill estimator for  $\alpha^{uu}$  under unanimous underreporting

by equation (3.19) thus reads

$$\hat{\alpha}^{uu}(w_{min}^{uu}; N) = N \cdot \left( \sum_{s=1}^N \ln \left( \frac{w^{uu}(s)}{w_{min}^{uu}} \right) \right)^{-1} \quad (3.20)$$

$$= N \cdot \left( \sum_{s=1}^N \ln \left( \frac{\rho \cdot w_{min} \cdot N}{\rho \cdot w_{min} \cdot s} \right) \right)^{-1} \quad (3.21)$$

$$= N \cdot \left( \sum_{s=1}^N \ln \left( \frac{N}{s} \right) \right)^{-1} \quad (3.22)$$

which is the unbiased estimator of equation (3.17).

### Differential Non-Response of the Upper $q$ Quantile. Asymptotic Properties.

Suppose that a discrete quantity  $w$  is distributed according to Zipf's law but  $q$ -truncated during measurement. The  $q$ -truncated rank-size rule therefore reads

$$w(s) = \frac{w_{min} \cdot N}{s}, \quad (3.23)$$

now with  $s = \lfloor q \cdot N \rfloor + 1, \lfloor q \cdot N \rfloor + 2, \dots, N$  as the ranks. For the  $q$ -truncated distribution, the MLE for the tail index under differential non-response, denoted  $\hat{\alpha}^{nr}$ , becomes

$$\hat{\alpha}^{nr}(q; N) = (N - \lfloor q \cdot N \rfloor + 1) \cdot \left( \sum_{s=\lfloor q \cdot N \rfloor + 1}^N \ln \left( \frac{N}{s} \right) \right)^{-1}. \quad (3.24)$$

Further simplifying equation (3.24) yields

$$\hat{\alpha}^{nr}(q; N) = \frac{1 + N - N \cdot q}{\ln \frac{N^{(N-N \cdot q)} \cdot (Nq)!}{N!}}. \quad (3.25)$$

Utilizing Stirling's approximation, in particular Ramanujan's version  $\ln n! \approx n \cdot \ln n - n + \frac{1}{6} \ln(n(1 + 4n(1 + 2n))) + \frac{1}{2} \pi$  (Ramanujan 1988), equation (3.25) now becomes

$$\hat{\alpha}^{nr}(q; N) \approx \frac{1 + N - N \cdot q}{u}. \quad (3.26)$$

with  $u = (N - N \cdot q) \cdot \ln(N) - \left[ N \cdot \ln N - N + \frac{1}{6} \ln(N(1 + 4N(1 + 2N))) \right] + \left[ N \cdot q \cdot \ln(N \cdot q) - N \cdot q + \frac{1}{6} \ln(N \cdot q(1 + 4(N \cdot q)(1 + 2 \cdot (N \cdot q)))) \right]$ .<sup>‡</sup> Finally, taking the limit of expression (3.26) yields

$$\lim_{N \rightarrow \infty} \frac{1 + N - N \cdot q}{u} = \frac{1 - q}{1 - q + q \cdot \ln(q)}. \quad (3.27)$$

---

<sup>‡</sup>In particular, (Ramanujan 1988) shows that the asymptotic error for the above approximation is  $\frac{1}{1440 N^3}$ , which suffices for the current purpose.



In the limit, the impact of  $N$  has completely vanished and the distortion of  $\alpha$  is now only dependent on  $q$ . As we easily see, even for large values of  $N$ , the estimator is (upward) biased for any positive  $q$ , since for any  $q > 0$ , the numerator is larger than the denominator, so  $\alpha > 1$ . The result in equation (3.27) shows that the upward bias is not merely an artefact of sample size, but holds true for any sufficiently large  $N$ .

**Differential Underreporting of the Upper  $q$  Quantile.** Asymptotic Properties.

Suppose that a quantity  $w$  is distributed according to Zipf's law. Consider the case, where only the upper  $q$ -quantile is proportionally underreporting with rate  $\rho \in (0, 1)$ . The rank-size rule is now a piecewise function for the upper  $q$ -quantile and the remaining  $1 - q$ , yielding

$$w(s)^{du} = \rho \cdot \frac{w_{min} \cdot N}{s}, \quad (3.28)$$

with  $s = 1, 2, \dots, \lfloor q \cdot N \rfloor$  as the ranks and

$$w(s)^{du} = \frac{w_{min} \cdot N}{s} \quad (3.29)$$

with  $s = \lfloor q \cdot N \rfloor + 1, \lfloor q \cdot N \rfloor + 2, \dots, N$  as the remaining ranks.

We require that  $w_{min}$ , the minimum of the unchanged initial distribution stays the minimum for the distribution with differential underreporting to avoid issues with the MLE which is based on this minimum. For this, the smallest reported value in the underreporting region has to be greater than  $w_{min}$ , that is,

$$w(qN)^{du} = \rho \cdot \frac{w_{min} \cdot N}{qN} > w_{min}$$

and therefore

$$\frac{\rho}{q} > 1. \quad (*)$$

Thus, for the minimum not to be affected, it has to hold by (\*) that the reporting rate exceeds the affected population share of the highest wealth levels. By the linearity of the sum function and assuming condition (\*) to hold, we obtain the Hill estimator for the tail exponent  $\hat{\alpha}^{du}$  under differential (proportional) underreporting as

$$\hat{\alpha}^{du}(q; \rho, N) = N \left( \sum_{s=\lfloor q \cdot N \rfloor + 1}^N \ln \left( \frac{N}{s} \right)^{-1} + \sum_{s=1}^{\lfloor q \cdot N \rfloor} \ln \left( \frac{\rho \cdot N}{s} \right)^{-1} \right). \quad (3.30)$$

Further simplifying yields

$$\hat{\alpha}^{du}(q; \rho, N) = \frac{N}{\ln \left( \frac{N^{(N-Nq)}(Nq)!}{(N-Nq)!} \right) + \ln \left( \frac{(N\rho)^{Nq}}{(Nq)!} \right)}. \quad (3.31)$$

Utilizing again Stirling's approximation, we get

$$\hat{\alpha}^{du}(q; \rho, N) \approx \frac{N}{v}, \quad (3.32)$$

with  $v = -\frac{1}{6}\ln \left( 8N^3 + 4N^2 + N + \frac{1}{30} \right) + Nq \ln(N\rho) + (N-Nq) \ln(N) + N + N(-\ln(N)) - \frac{\ln(\pi)}{2}$ .

Taking the limit for  $N \rightarrow \infty$  gives

$$\lim_{N \rightarrow \infty} \frac{N}{v} = \frac{1}{1 + q \ln(\rho)}. \quad (3.33)$$

Notice that for  $q \in (0, 1)$  and  $\rho \in (q, 1)$ , the estimator is therefore always upward biased compared to the Zipf benchmark of  $\alpha = 1$ . Condition (\*) precludes the possibility of a negative induced bias which would result from  $q \ln(\rho) < -1$  and would be uninterpretable. For this, note that condition (\*) implies  $\ln(\rho) > \ln(q)$ , since  $\ln(\cdot)$  is monotonically increasing in its argument. It is thus sufficient to show that  $q \ln(q) > -1$ . Rearranging yields

$$\ln(q) + \frac{1}{q} > 0. \quad (3.34)$$

Define  $f(q) = \ln(q) + \frac{1}{q}$ . By

$$f(1) = 1 \quad (3.35)$$

$$\text{and } \frac{df}{dq} = \frac{1}{q} - \frac{1}{q^2} < 0, \forall q \in (0, 1), \quad (3.36)$$

we know that the function is monotonically decreasing for the whole considered interval and positive at the upper interval boundary. From (3.35) and (3.36), we can therefore conclude that  $f(q) > 0$  for all  $q \in (0, 1)$ . This is exactly the non-negativity constraint in (3.34) and thus the desired result that condition (\*) implies strictly non-negative induced biases in approximation (3.33).

### 3.C Empirical Densities of Returns to Wealth Portfolios

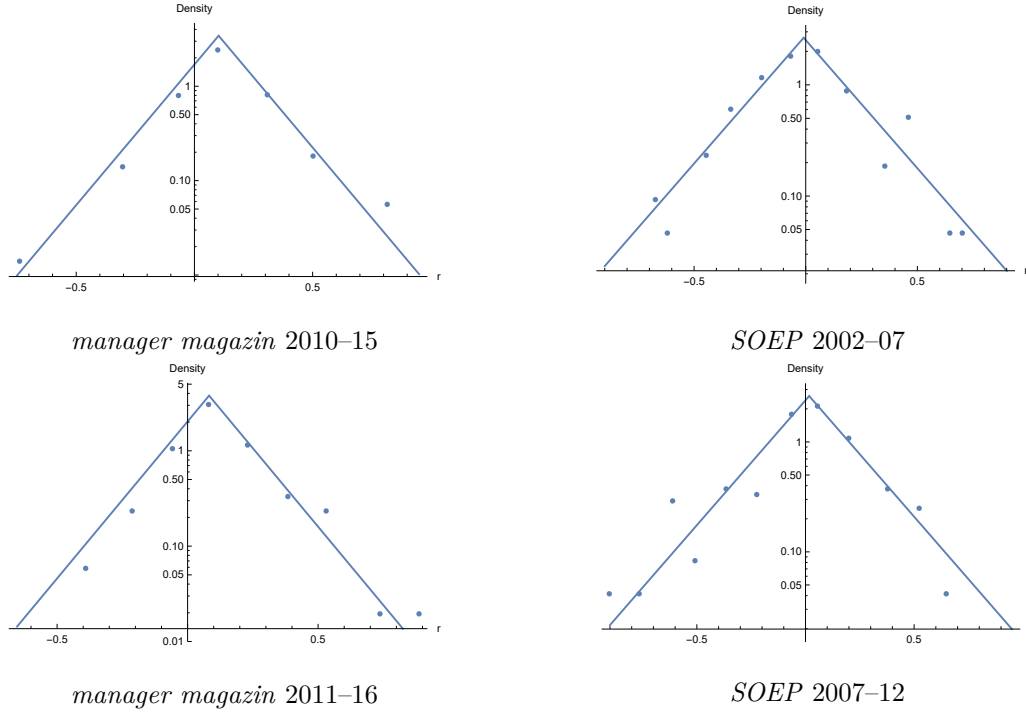


Figure 3.2: Empirical densities for the distribution of returns to wealth portfolios across different sample types and time periods, with maximum likelihood fits of the Laplace distribution indicated by solid lines.

Note that returns for the rich list samples exhibit a positive median, while the median for the survey samples is indistinguishable from zero at the usual significance levels. The reason we consider the median (instead of, say, the expectation or mode) is that the median corresponds to the maximum likelihood estimate of the location parameter for the Laplace distribution (3.4). By the same token, the MLE of the dispersion parameter in (3.4) is the mean absolute deviation, not the variance. Overall, visual inspection already confirms the Laplacian nature of wealth returns, since the empirical densities exhibit a linear tent shape on semi-logarithmic scale that is characteristic of the Laplace distribution. To test parametrically for the hypothesis of a Laplace distribution in wealth returns, we follow standard procedure and consider the exponential power (or Subbotin) distribution, because the Laplace is a special case of the Subbotin when the shape parameter  $\kappa$  equals unity. This is readily verified from the Subbotin density given by

$$f(r, \kappa, \sigma, m) = \frac{1}{2\sigma\kappa^{1/\kappa}\Gamma(1 + 1/\kappa)} \exp\left(-\frac{1}{\kappa} \left|\frac{r - m}{\sigma}\right|^\kappa\right), \quad (3.37)$$

where  $\kappa, \sigma \in \mathbb{R}^+, m \in \mathbb{R}$ , and  $\Gamma(\cdot)$  denotes the Gamma function. MLEs of the shape parameter are reported in the table below, showing that we cannot reject the Laplace hypothesis in our data.

	Subbotin shape parameter $\hat{\kappa}$	Standard error
<i>SOEP</i>		
2002–07	1.18	0.1766
2007–12	0.8463	0.1178
<i>manager magazin</i>		
2010–15	0.8523	0.08587
2011–16	0.9509	0.09798

Table 3.7: Maximum likelihood estimates of the Subbotin shape parameter, denoted  $\hat{\kappa}$ , for the distribution of wealth returns cannot reject the Laplace hypothesis at the usual significance levels. We employed *Subbotools* 1.3.0 for estimation as it delivers the most accurate and efficient estimates in simulation runs (Bottazzi 2014).

### 3.D Total Wealth Estimates

We estimate the total wealth levels  $\hat{W}$  by numerical integration according to

$$\hat{W}_i = \hat{n}^i \cdot \int_{\hat{w}_{min}^{SOEP}}^{w_{max}^{mm}} \hat{f}_i(w) \cdot w \, dw, \quad (3.38)$$

where  $\hat{f}_i(w)$  denotes the estimated PDF of the power law given by

$$\hat{f}_i(w) = \hat{\alpha}^i \cdot \hat{w}_{min}^{SOEP} \cdot w^{-(\hat{\alpha}^i+1)}, \quad (3.39)$$

where  $\hat{w}_{min}^{SOEP}$  is the estimated minimum from the *SOEP* sample and  $w_{max}^{mm}$  is the maximum from the rich list, and are common across approaches, while the estimated population  $\hat{n}^i$ , and the estimated tail index  $\hat{\alpha}^i$  are chosen dependent on the pre-analytical vision. There are three possible  $\hat{\alpha}^i \in \{\hat{\alpha}^{SOEP}; \hat{\alpha}^{mm}; \hat{\alpha}^{Zipf} = 1\}$ , corresponding to the estimates from the survey study, the rich lists and for Zipf’s law, respectively. The same holds for the estimated population.

The population estimates based on the *SOEP* extrapolate the in-sample ratio of the power law population to the whole German population. In particular, let  $\omega$  be the ratio of the estimated power law population relative to the whole sample size. For the uncorrected estimates, we set the sample size to 0.035% of 82,500,000 which equals 28,875, where 0.035% is the approximate sample ratio in the *SOEP* surveys. The power law population in Germany is then calculated as  $\omega \cdot N$  for the uncorrected case  $\hat{n}^{SOEP}$ , with  $N = 82,500,000$  (Statistisches Bundesamt 2017). For the Zipf case, we correct  $\omega$  by the estimated non-response rates in Table 3.5 and calculate  $\hat{n}^{Zipf} = N(\omega + q_{tot})$ . Given the relatively

low estimates of  $q_{tot}$ , the estimated population levels do not differ too much from the uncorrected estimates. The results are reported in Table 3.8.

Population from <i>SOEP</i>		
	Uncorrected	Corrected
Population estimate $\hat{n}$ (2002)	2,745,714	2,753,993
Population estimate $\hat{n}$ (2007)	3,600,000	3,603,629
Population estimate $\hat{n}$ (2012)	3,805,714	3,820,703

Table 3.8: The estimates assume a total population  $N = 82,500,000$  and are calculated from the in-sample power law population fractions  $\omega$ , and the estimated non-response rates  $q_{pl}$  for the corrected case.

We estimate the different population levels for the *manager magazin* by taking the CDF  $P(w; w_{min}, \alpha)$  of a continuous power law with parameters  $\hat{w}_{min}^{SOEP}$  from the *SOEP* samples and the tail indices  $\hat{\alpha}^{mm}$ , the in-sample power law populations  $N^{mm}$  and the minima  $w_{min}^{mm}$  from the *manager magazin* samples. For a specific parameter combination,  $\hat{n}^{mm}$  is then calculated as  $\hat{n}^{mm} = 1/(1 - P(w_{min}^{mm}; \hat{w}_{min}^{SOEP}, \hat{\alpha}^{mm}) \cdot N^{mm})$ . The intuition is that  $\hat{n}^{mm}$  corresponds to the power law population when the power law in the *manager magazin* sample is extended to the minima determined from the *SOEP* surveys. The results are reported in Table 3.9 below.

Population from <i>mm</i>						
	2010	2011	2013	2014	2015	2016
Population estimate $\hat{n}$ ( $w_{min} = 280,000$ )	352,534	324,983	869,368	407,397	288,683	88,307
Population estimate $\hat{n}$ ( $w_{min} = 200,000$ )	493,270	452,876	1,258,710	567,946	395,524	114,097
Population estimate $\hat{n}$ ( $w_{min} = 180,000$ )	547,981	502,467	1,413,360	630,216	436,510	123,628

Table 3.9: The estimates are calculated from the parameter combinations of the various estimated minima  $\hat{w}_{min}^{SOEP}$  in the *SOEP* samples, and the tail indices  $\hat{\alpha}^{mm}$ , the in-sample power law populations  $N^{mm}$ , and the minima  $w_{min}^{mm}$  of the *manager magazin* samples for each of the respective years given in the Table.

### 3.E Equiprobable Sampling From a Power Law

**Simple Random Sampling without Replacement.** Equiprobable Selection of Elements.

Let  $\mathcal{N}$  denote the total population with size  $N \in \mathbb{N}^+$  and  $\mathcal{S}$  a sample out of  $\mathcal{N}$  with size  $S \in \mathbb{N}^+$  and  $S \leq N$ . The sampling procedure selects each element of the set  $\mathcal{N}$  with equal probability and without replacement. If the maximum value of  $\mathcal{N}$ , denoted by  $w_{max}$ , is unique, the probability of  $w_{max}$  to be included in the sample  $\mathcal{S}$ , that is,  $p(w_{max} \in \mathcal{S})$ , is equivalent to the probability of any unique element to be chosen under these conditions. The inclusion probability of  $w_{max}$  in the chosen set  $\mathcal{S}$  is therefore given by

$$p(w_{max} \in \mathcal{S}) = \frac{\binom{N-1}{S-1}}{\binom{N}{S}} \quad (3.40)$$

$$= \frac{(N-1)!}{(S-1)! \cdot (N-S)!} \cdot \frac{N!}{N!} \quad (3.41)$$

$$= \frac{S! \cdot (N-S)!}{N! \cdot (S-1)!} \quad (3.42)$$

$$= \frac{S \cdot (S-1)! \cdot (N-1)!}{N \cdot (S-1)! \cdot (N-1)!} \quad (3.43)$$

$$= \frac{S}{N}. \quad (3.44)$$

The inclusion probability under simple random sampling without replacement for  $w_{max}$  therefore corresponds to the sampling ratio  $S/N$  and is equal to unity only if  $S = N$ .

### 3.F Logarithmic Sampling From a Power Law

**Logarithmic Random Sampling without Replacement.** Assumption of Zipf's Law.

Let again  $\mathcal{N}$  with size  $N \in \mathbb{N}^+$  denote the total population and  $\mathcal{S}$  the sample out of  $\mathcal{N}$  with size  $S \in \mathbb{N}^+$  and  $S \leq N$ . Furthermore, assume that the total population is now divided into  $v$  different intervals or “slices”, where the length of each slice corresponds to one order of magnitude of the relevant quantity  $w$ , so the intervals are scaled logarithmically. Now the same procedure as above is applied to each logarithmic slice, that is every element in each slice is selected with equal probability and without replacement. It has to hold that  $v \in \mathbb{N}^+$ .

For every slice,  $S/v$  elements are included in the sample of size  $S$ , where  $S$  obviously needs to be an integer multiple of  $v$ , since  $(S/v) \in \mathbb{N}^+$ . The slice covering the highest order of magnitude for  $w$  also has to include  $w_{max}$  as the maximum value. If one assumes Zipf's law to hold, this range of  $w$  includes a proportion  $10^{-v+1}$  of the total population with size  $N$ . Under Zipf's law this procedure chooses  $S/v$  elements out of a set of size

$N/10^{v-1}$ . The probability to include  $w_{max}$  in the chosen set  $\mathcal{S}$  is therefore

$$p(w_{max} \in \mathcal{S}) = \frac{\binom{(N/10^{v-1})-1}{(S/v)-1}}{\binom{N/10^{v-1}}{S/v}} \quad (3.45)$$

$$= \frac{((N/10^{v-1}) - 1)!}{((S/v) - 1)! \cdot ((N/10^{v-1}) - (S/v))!} \quad (3.46)$$

$$= \frac{(S/v)! \cdot ((N/10^{v-1}) - (S/v))!}{(N/10^{v-1})! \cdot ((S/v) - 1)!} \quad (3.47)$$

$$= \frac{(S/v) \cdot ((S/v) - 1)! \cdot ((N/10^{v-1}) - 1)!}{(N/10^{v-1}) \cdot ((N/10^{v-1}) - 1)! \cdot ((S/v) - 1)!} \quad (3.48)$$

$$= \frac{10^{v-1}}{v} \cdot \frac{S}{N}, \quad \text{with } \frac{S}{N} \geq \frac{10^v}{v}.$$

Therefore, the inclusion probability  $p(w_{max} \in \mathcal{S})$  under logarithmic sampling converges  $10^{v-1}/v$  times faster to unity compared to simple random sampling. The sampling ratio  $S/N$  has to equal merely  $v/10^{v-1}$  for  $p(w_{max} \in \mathcal{S}) = 1$ . For  $v = 2$ , it has to equal 0.2, for  $v = 3$ , it has to equal 0.03, and so on. Of course, this is the case because every element in the interval covering the highest order of magnitude for  $w$  has to be included in the sample. For  $v = 1$ , this procedure obviously corresponds to the case of pure equiprobable sampling, where the inclusion probability is equal to the sampling ratio  $S/N$ , as  $(10^{v-1}/v) = 1$  for  $v = 1$ .

## References

- Alan, S., Atalay, K., and Crossley, T. F. (2015). Do the rich save more? Evidence from Canada. *Review of Income and Wealth*, 61(4):739–758.
- Alstadsæter, A., Johannesen, N., and Zucman, G. (2019). Tax evasion and inequality. *American Economic Review*, 109(6):2073–2103.
- Andersen, T. G., Bollerslev, T., Christoffersen, P. F., and Diebold, F. X. (2006). Volatility and correlation forecasting. In Granger, C. W. J., Elliott, G., and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 1, pages 777–878. Elsevier.
- Aoki, S. and Nirei, M. (2016). Pareto distribution of income in neoclassical growth models. *Review of Economic Dynamics*, 2:25–42.
- Bach, L., Calvet, L. E., and Sodini, P. (2017). Rich pickings? Risk, return, and skill in the portfolios of the wealthy. *HEC Paris Research Paper*, 1126:16–03.

- Bach, S., Thiemann, A., and Zucco, A. (2019). Looking for the missing rich: Tracing the top tail of the wealth distribution. *International Tax and Public Finance*, 26(6):1234–1258.
- Bali, T. G. (2008). The intertemporal relation between expected return and risk. *Journal of Financial Economics*, 8:101–131.
- Bali, T. G. and Peng, L. (2006). Is there a risk-return trade-off? Evidence from high-frequency data. *Journal of Applied Econometrics*, 21(8):1169–1198.
- Balz, J., Sunstein, C., and Thaler, R. (2014). Choice architecture. In Shafir, E., editor, *The behavioral foundations of public policy*, pages 428–439. Princeton University Press, Princeton.
- Benhabib, J. and Bisin, A. (2018). Skewed wealth distributions: Theory and empirics. *Journal of Economic Literature*, 56(4):1261–91.
- Benhabib, J., Bisin, A., and Zhu, S. (2011). The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1):123–175.
- Blank, A. and Solomon, S. (2000). Power laws in cities population, financial markets and internet sites (scaling in systems with a variable number of components). *Physica A: Statistical Mechanics and its Applications*, 287(1-2):279–288.
- Bottazzi, G. (2014). Subbotools user’s manual. *Laboratory of Economics and Management Working Paper Series, Sant’Anna School of Advanced Studies, Pisa (Italy)*, March 2014.
- Brandt, M. W. and Kang, Q. (2004). On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach. *Journal of Financial Economics*, 72(2).
- Calvet, L. E. and Sodini, P. (2014). Twin picks: Disentangling the determinants of risk-taking in household portfolios. *Journal of Finance*, 69(2):867–906.
- Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics*, 18(2):373–399.
- Campbell, J. Y. and Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, 31(3):281–318.
- Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 5:661–703.
- Córdoba, J. C. (2008a). A generalized Gibrat’s law. *International Economic Review*, 49(4):1463–1468.
- Córdoba, J.-C. (2008b). On the distribution of city sizes. *Journal of Urban Economics*, 63(1):177–197.
- Cristelli, M., Batty, M., and Pietronero, L. (2012). There is more than a power law in Zipf. *Scientific Reports*, 2:812.



- Davies, J. B. and Shorrocks, A. F. (2000). The distribution of wealth. In Atkinson, A. and Bourguignon, F., editors, *Handbook of Income Distribution*, pages 605–675. Elsevier Science B.V., Amsterdam.
- De Haan, L. and Resnick, S. (1997). On asymptotic normality of the Hill estimator. *Communications in Statistics. Stochastic Models*, 14(4):849–866.
- Deaton, A. (2003). Saving and liquidity constraints. *Econometrica*, 59(5):1221–1248.
- Dynan, K. E., Skinner, J., and Zeldes, S. P. (2004). Do the rich save more? *Journal of Political Economy*, 112(2):397–444.
- Eckerstorfer, P., Halak, J., Kapeller, J., Schütz, B., Springholz, F., and Wildauer, R. (2016). Correcting for the missing rich: An application to wealth survey data. *Review of Income and Wealth*, 62(4):605–627.
- Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2020). Heterogeneity and persistence in returns to wealth. *Econometrica*, 88(1):115–170.
- Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business*, 38(1):34–105.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417.
- Fama, E. F. (1991). Efficient capital markets: II. *Journal of Finance*, 46(5):1575–1617.
- French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19(1):3–29.
- Frick, J. R., Grabka, M. M., and Marcus, J. (2007). Editing and multiple imputation of item-non-response in the 2002 wealth module of the German Socio-Economic Panel (SOEP). *DIW Berlin: Data Documentation*, 18.
- Gabaix, X. (1999). Zipf’s law for cities: An explanation. *Quarterly Journal of Economics*, 114(3):739–767.
- Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics*, 1:255–293.
- Gabaix, X., Lasry, J.-M., Lions, P.-L., and Moll, B. (2016). The dynamics of inequality. *Econometrica*, 84(6):2071–2111.
- Galbraith, J. K. (2019). Sparse, inconsistent and unreliable: Tax records and the World Inequality Report 2018. *Development and Change*, 50(2):329–346.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, 76(3):509–548.
- Gibrat, R. (1931). *Les inégalités économiques; applications: aux inégalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d’une loi nouvelle, la loi de l’effet proportionnel*. Librairie du Recueil Sirey.

- Glazer, E. (2016). At J.P. Morgan, \$9 million in assets isn't rich enough. *Wall Street Journal*, March, 18.
- Goldstein, M. L., Morris, S. A., and Yen, G. G. (2004). Problems with fitting to the power-law distribution. *European Physical Journal B*, 41(2):255–258.
- Gonzales, M., Nave, J., and Rubio, G. (2012). The cross section of expected returns with MIDAS betas. *Journal of Financial and Quantitative Analysis*, 47(1):115–135.
- Goyal, A. and Santa-Clara, P. (2003). Idiosyncratic risk matters! *Journal of Finance*, 58(3):975–1007.
- Guiso, L., Jappelli, T., and Terlizzese, D. (1996). Income risk, borrowing constraints, and portfolio choice. *American Economic Review*, 86(1):158–172.
- Guo, H. and Whitelaw, R. F. (2006). Uncovering the risk–return relation in the stock market. *Journal of Finance*, 61(3):1433–1463.
- Harvey, C. R. (2001). The specification of conditional expectations. *Journal of Empirical Finance*, 8(5):573–637.
- Heitjan, D. F. and Rubin, D. B. (1991). Ignorability and coarse data. *Annals of Statistics*, 19(4):2244–2253.
- Hubmer, J., Krusell, P., and Smith Jr, A. A. (2016). The historical evolution of the wealth distribution: A quantitative-theoretic investigation. *NBER Working Paper Series*, 23011.
- Jappelli, T. and Pistaferri, L. (2014). Fiscal policy and MPC heterogeneity. *Macroeconomics*, 6(4):107–136.
- Jones, C. I. and Kim, J. (2018). A Schumpeterian model of top income inequality. *Journal of Political Economy*, 126(5):1785–1826.
- Kennickell, A. B. and Woodburn, R. L. (1999). Consistent weight design for the 1989, 1992 and 1995 SCFs, and the distribution of wealth. *Review of Income and Wealth*, 45(2):193–215.
- King, M. A. and Leape, J. I. (1998). Wealth and portfolio composition: Theory and evidence. *Journal of Public Economics*, 69(2):155–193.
- King, M. R. and Maier, P. (2010). Would greater regulation of hedge funds reduce systemic risk? In Kolb, R. W., editor, *Lessons from the Financial Crisis: Causes, Consequences, and Our Economic Future*, pages 625–632. John Wiley & Sons, Hoboken, New Jersey.
- Kotz, S., Kozubowski, T. J., and Podgorski, K. (2001). *The Laplace distribution and generalizations*. Birkhauser, Boston.
- Levy, M. (2003). Are rich people smarter? *Journal of Economic Theory*, 11:42–64.
- Levy, M. and Levy, H. (2003). Investment talent and the Pareto wealth distribution: Theoretical and experimental analysis. *Review of Economics and Statistics*, 85(3):709–725.

- Lundblad, C. (2007). The risk return tradeoff in the long run: 1836–2003. *Journal of Financial Economics*, 85(1):123–150.
- Luttmer, E. G. (2010). Models of growth and firm heterogeneity. *Annual Review of Economics*, 2(1):547–576.
- Luttmer, E. G. (2011). On the mechanics of firm growth. *Review of Economic Studies*, 78(3):1042–1068.
- Luttmer, E. G. (2018). Slow convergence in economies with organization capital. *University of Minnesota and Federal Reserve Bank of Minneapolis Working Paper Series*, 748.
- Malcai, O., Biham, O., and Solomon, S. (1999). Power-law distributions and Levy-stable intermittent fluctuations in stochastic systems of many autocatalytic elements. *Physical Review E*, 60(2):1299.
- Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–887.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370.
- Nirei, M. (2009). Pareto distributions in economic growth models. *Institute of Innovation Research Working Paper*, 09(05).
- Nirei, M. and Souma, W. (2007). A two factor model of income distribution dynamics. *Review of Income and Wealth*, 53(3):440–459.
- Piketty, T. and Zucman, G. (2015). Wealth and inheritance in the long run. In Atkinson, A. and Bourguignon, F., editors, *Handbook of Income Distribution*, volume 2, pages 1303–1368. Elsevier Science B.V., Amsterdam.
- Quadrini, V. (1999). The importance of entrepreneurship for wealth concentration and mobility. *Review of Income and Wealth*, 45(1):1–19.
- Quine, W. V. (1975). On empirically equivalent systems of the world. *Erkenntnis*, 9(3):313–328.
- Ramanujan, S. (1988). *The lost notebook and other unpublished papers*. Narosa Publishing House, New Delhi, Berlin, New York.
- Schneeweiß, H., Komlos, J., and Ahmad, A. S. (2010). Symmetric and asymmetric rounding: A review and some new results. *AStA Advances in Statistical Analysis*, 94(3):247–271.
- Schröder, C., Bartels, C., Göbler, K., Grabka, M. M., and König, J. (2020). MillionärInnen unter dem Mikroskop: Datenlücke bei sehr hohen Vermögen geschlossen – Konzentration höher als bisher ausgewiesen. *DIW Wochenbericht*, 29:511–522.
- Statistisches Bundesamt (2017). Bevölkerung in Deutschland voraussichtlich auf 82,8 Millionen gestiegen. *Pressemitteilung*, 33/17.

- Stumpf, M. P. H. and Porter, M. (2012). Critical truths about power laws. *Science*, 335(6069):665–666.
- Subbotin, M. T. (1923). On the law of frequency of error. *Matematicheskii Sbornik*, 31(2):296–301.
- Toda, A. A. (2014). Incomplete market dynamics and cross-sectional distributions. *Journal of Economic Theory*, 154:310–348.
- Vermeulen, P. (2016). Estimating the top tail of the wealth distribution. *American Economic Review*, 106(5):646–50.
- Vermeulen, P. (2018). How fat is the top tail of the wealth distribution? *Review of Income and Wealth*, 64(2):357–387.
- Wachter, J. A. and Yogo, M. (2010). Why do household portfolio shares rise in wealth? *Review of Financial Studies*, 23(11):3929–3965.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley Press, Boston.

# Equal Chances, Unequal Outcomes? Network-Based Evolutionary Learning and the Industrial Dynamics of Superstar Firms

---

Joint work with D. M. Mayerhoffer.

Published under Creative Commons Attribution 4.0 License:

Schulz, J. and Mayerhoffer, D. M. (2021). Equal chances, unequal outcomes? Network-based evolutionary learning and the industrial dynamics of superstar firms. *Journal of Business Economics*, 91(9):1357–1385. <https://doi.org/10.1007/s11573-021-01047-8>

The authors thank Giovanni Dosi and Marcello Pereira for sharing their code and explaining their rationale; for helpful comments on style and content (in alphabetical order) Anna Gebhard, Bettina Gregg, Torsten Heinrich, Dominik Klein, Johannes Marx, Alessandro Lucini Paioni, Roland Parr, Simon Scheller, Jürgen Sirsch, Moritz Schulz and Jan Weber as well as the participants of the WICK 2020 and EAEPE 2020 conferences. We are especially indebted to two anonymous reviewers and the editor Tim Weitzel for multiple insightful comments and suggestions to improve our paper. All remaining errors are our own, of course. This paper first circulated under the title “Innovation of Incumbents and Entrants in Localised Competition”. We are immensely grateful to Tim Weitzel for pointing us to a more appropriate title.

---

## Abstract

With the advent of platform economies and the increasing availability of online price comparisons, many empirical markets now select on relative rather than absolute performance. This feature might give rise to the ‘winner takes all/most’ phenomenon, where tiny initial productivity differences amount to large differences in market shares. We study the effect of heterogeneous initial productivities arising from locally segregated markets on aggregate outcomes, e.g., regarding revenue distributions. Several of those firm-level characteristics follow distributional regularities or ‘scaling laws’ (Brock 1999). Among the most prominent are Zipf’s law describing the largest firms’ extremely concentrated size distribution and the robustly fat-tailed nature of firm size growth rates, indicating a high frequency of extreme growth events. Dosi et al. (2017b) recently proposed a model of evolutionary learning that can simultaneously explain many of these regularities. We propose a parsimonious extension to their model to examine the effect for deviations in market structure from global competition, implicitly assumed in Dosi et al. (2017b). This extension makes it possible to disentangle the effects of two modes of competition: the global competition for sales and the localised competition for market power, giving rise to industry-specific entry productivity. We find that the empirically well-established combination of ‘superstar firms’ and Zipf tail is consistent only with a knife-edge scenario in the neighbourhood of most intensive local competition. Our model also contests the conventional wisdom derived from a general equilibrium setting that maximum competition leads to minimum concentration of revenue (Silvestre 1993). We find that most intensive local competition leads to the highest concentration, whilst the lowest concentration appears for a mild degree of (local) oligopoly. Paradoxically, a level playing field in initial conditions might induce extreme concentration in market outcomes.

---

---

## Contents of this Paper

4.1	Introduction . . . . .	87
4.2	Models of Evolutionary Learning as a Representation of Empirical Findings	89
4.2.1	Stylised Facts in Industrial Dynamics . . . . .	89
4.2.2	Agent-Based Models . . . . .	91
4.2.3	An Agent-Based Model of Learning and Selection . . . . .	91
4.2.4	Productivity Gain Through Stochastic Learning . . . . .	92
4.3	Model Description . . . . .	94
4.3.1	Model Properties and Initialisation . . . . .	94
4.3.2	Events During the Simulation . . . . .	96
4.4	Simulation Results . . . . .	98
4.4.1	Size Distribution . . . . .	99
4.4.2	Growth Rates of Market Shares . . . . .	102
4.4.3	Age . . . . .	105
4.4.4	Generating Mechanisms . . . . .	106
4.5	Discussion . . . . .	108
	References . . . . .	112

---

## 4.1 Introduction

Within an increasing number of markets, an individual firm's fate is no longer determined by absolute performance but by its performance relative to its competitors. Put differently, individual success in the market is a function of the now widespread availability of price comparisons on the internet (Akerman et al. 2022) and platform competition (Autor et al. 2020). In those 'winner takes all/most' markets, tiny differences in initial productivity can manifest themselves into large differences in market shares, typically leading to a high emergent concentration of market power. Somewhat surprisingly, the determinants of initial productivity at market entry have received little scholarly attention. We systematically explore the aggregate effects of heterogeneous initial conditions by exploiting a plausible notion of industry-specific productivity within locally segregated markets. Our approach builds on the intuition that comparisons of relative performance are seldom global and typically localised. For example, Uber and Alphabet not competing within the same submarket, even though their business strategies building on network effects and intangibles appear to be very similar. Our results suggest that market segmentation exhibits sizeable and counter-intuitive effects on the distribution of market shares, on firm growth and firm survival.

Empirically, these aforementioned firm-level characteristics follow distributional regularities or 'scaling laws' (Brock 1999), whose underlying mechanisms require expation. Dosi

et al. (2017b) recently proposed a model of industrial dynamics that features evolutionary learning: Individual firms innovate and increase their productivity but compete for market shares according to a global selection mechanism based on productivity. This *evolutionary learning* mechanism combines cumulative learning with a ‘winner takes all/most’ market structure. Despite its bare-bones, partial equilibrium nature, the model is able to explain a surprising number of stylised facts in industrial dynamics, such as strongly heterogeneous size distributions, scaling between size growth rates and their variance as well as (persistent) heterogeneity in productivity. The model has been applied successfully and essentially unchanged in a macroeconomic setting both for explanatory purposes and policy experiments in the ‘Keynes meets Schumpeter’ (K+S) modelling approach (Dosi et al. 2010, 2013, 2015, 2017a). Apart from the K+S approach, the distributional regularities, which even the partial model produces, have been identified to be of great macroeconomic relevance (Gabaix 2011; Di Giovanni et al. 2011).

We propose a parsimonious extension to this model to examine the effect for deviations in market structure from global competition. Namely, we introduce a network structure of localised competition and innovation. This extension makes it possible to disentangle effects from two modes of competition: global competition for sales and localised competition for market power, giving rise to industry-specific productivity differences. Our contribution is thus twofold: Firstly, we test the benchmark model results’ robustness for different market structures, as defined by local competition for market shares and localised entry, where the entry process’ precise nature has recently been identified as the most important driver of aggregate outcomes within the model (Dosi et al. 2018). By this, we are able to constrain the range of possible competitive mechanisms in light of the empirical evidence in more detail. Secondly, we take a complementary approach to the macroeconomic implementations and examine even further the micro processes that different competition structures induce. These microeconomic considerations have consequences for regulatory policy and implications even at the managerial level, in particular, for market entry timings.

Besides the cumulative idiosyncratic learning mechanism from our benchmark, for a non-complete network, our model also features a second process that distinguishes between firms’ different productivity levels: When a new firm enters the model market, it acquires an industry-specific productivity level, modelled as the weighted average productivity of its direct competitors. This mechanism applies to a firm only once in its lifetime, namely at foundation/market entry. Nonetheless, it can crucially shape the entire life of a firm, since initial productivity determines whether an entrant can stabilise its position in the market or is quickly forced out of it again: We show that successful entrants typically join highly productive product markets. Hence, from a management perspective, our findings underline the importance of timing market entrance and thorough search prior to entry.



Namely, our model can explain, why Schlichte et al. (2019) find the most successful firms to be fast followers in innovative markets rather than the original innovators themselves. We find that the empirically well-established combination of ‘superstar firms’ and Zipf tails is consistent only with a knife-edge scenario in the neighbourhood of most intensive local competition. Moreover, our contests the conventional wisdom derived from a general equilibrium setting that maximum competition leads to minimum concentration of revenue (Silvestre 1993). Instead, we find that most intensive local competition leads to the highest concentration and the lowest concentration appears for a mild degree of (local) oligopoly. Relating to a different notion of competition, this finding might also be interpreted as evidence that ‘winner takes all/most’ markets are far from the ordoliberal ideal which considers “competition [to be] the most ingenious disempowerment instrument in history” (Böhm 1960, p. 22, author’s translation). By contrast, it is precisely the ordoliberal demands for a ‘level playing field’ in combination with ‘performance based competition’ which lead to the highest concentration of revenue and hence power asymmetries within such markets that ordoliberals hope to avoid (Dold and Krieger 2019).

The remainder of this paper is organised as follows: We firstly discuss the stylised empirical facts that we intend to study as well as concepts and models on which we build (Section 4.2), followed by a detailed introduction of our model, based on Dosi et al. (2017b) (Section 4.3). Thereafter, we present the core simulation outcomes and explain their generating mechanisms (Section 4.4). Finally, we situate our findings in economic, policy and business discussions, drawing practical as well as normative implications, before closing with proposals for further research Section 4.5).

## **4.2 Models of Evolutionary Learning as a Representation of Empirical Findings**

### **4.2.1 Stylised Facts in Industrial Dynamics**

As a selection criterion for the parameter range that our proposed model spans, we use a set of stylised facts from traditional microeconomic literature and the industrial dynamics literature on distributional regularities in various firm-specific variables. The most prominent of these regularities is the finding of Zipf’s law, originally based in linguistics (Zipf 1949), for the upper tail of firm size distributions. This implies that the size distributions of the largest firms are extremely concentrated, where the second-largest firm has only approximately one half of the size of the largest, the third-largest only a third and so on. Zipf’s law in firm sizes appears to be a genuine and universal characteristic of market economies. A non-exhaustive list of studies on Zipf’s findings include Axtell (2001) for the US, Di Giovanni et al. (2011) for France, Pascoal et al. (2016) for Portugal,

Okuyama et al. (1999) for Japan, Kang et al. (2011) for the Republic of Korea, Zhang et al. (2009) along with Heinrich and Dai (2016) for China, and Fujiwara et al. (2004) for several European countries. This empirical regularity not only constrains the set of possible generating mechanisms, but Zipf's law has also been linked to several important economic phenomena, such as the surge in CEO payments in recent decades (Gabaix and Landier 2008), the explanation of aggregate fluctuations from the micro-level together with increases in aggregate volatility (Gabaix 2011) and the welfare effects of barriers to entry and trade liberalisation (Di Giovanni and Levchenko 2013). This whole strand of macroeconomic literature takes Zipf's law as their starting point but does not examine possible conditions for which it emerges. Our findings on the determinants for Zipf's law might also provide insights on how to influence this wide range of phenomena, from CEO payments to aggregate fluctuations and international trade, which from a structural perspective this literature takes as given.

Another phenomenon of similar attributed economic relevance is the recent emergence of 'superstar firms', that operate in 'winner takes most/all' markets. They have experienced substantial and sustained increases in revenue over relatively short amounts of time (Autor et al. 2020). Anecdotal examples for this behaviour are Alphabet and Uber. The rise of these firms has been proposed as an explanation for the recent decline in the labour share of national income (Autor et al. 2020), otherwise famously staying constant throughout the most part of recorded history of capitalist economies (Kaldor 1961) and the rise of wage inequality (Gabaix and Landier 2008).

Growth in various measures of size such as gross sales, total assets or number of employees for the whole range of firms has also been shown to be fat-tailed with relatively frequent extreme events, where empirical growth rate densities display a characteristic 'tent shape' on a semi-logarithmic scale, implying an exponential power functional form (Amaral et al. 1997; Bottazzi et al. 2001, 2002; Bottazzi and Secchi 2005, 2006; Alfarano and Milaković 2008; Bottazzi et al. 2011; Erlingsson et al. 2013; Mundt et al. 2015). These distributions have frequently been identified as Laplacian (Kotz et al. 2012), which though has recently been challenged theoretically and empirically (Mundt et al. 2015). We stick to the exponential power or Subbotin (1923) shape and focus on the fat-tailed nature of size growth rates.

Concerning the firm age distribution, empirical findings are scarcer. However, the (limited) consensus appears to be that firm age is approximated well by an exponential distribution, as shown by a number of studies: Coad and Tamvada (2008) for several developing countries; Kinsella (2009) for Irish firms; Coad (2010) for the plant level in the US and Daepf et al. (2015) for publicly listed firms in the US. This exponential age distribution also has the crucial theoretical implication that death rates are independent of firm age, as Daepf et al. (2015) confirm empirically.

The main stylised facts which our model aims to replicate are thus: a Zipf law in the upper tail of firm sizes;<sup>1</sup> fat-tailed growth behaviour; and a high frequency of ‘superstar’ high growth events coupled with an exponential age distribution with a common insolvency probability for all firms, irrespective of age.

## 4.2.2 Agent-Based Models

Competition within industries, which produces the stylised facts discussed here, constitutes a socio-economic system. To study the dynamics in this system and find candidates for mechanisms that lead to empirically observed facts, an Agent-Based Model (ABM) forms an adequate approach (Klein et al. 2018): Due to competitive interactions between individual firms, one cannot properly describe the system by additive aggregation of the model, but observed phenomena are emergent (Coleman 1990). Agent-Based Models can highlight and explain emergent phenomena and open the black box of competitive interactions on a system level in order to uncover interactions in and structures of subsystems (Hedström and Ylikoski 2010). Given the high level of idealisation, we do not intend to make quantitative predictions, but are nevertheless confident that the model can reveal central qualitative features of its economic target system (Grüne-Yanoff 2009).

## 4.2.3 An Agent-Based Model of Learning and Selection

We introduce a layer of locality to the model of learning and selection of industrial dynamics by Dosi et al. (2017b).<sup>2</sup> This model understands learning as an increase in productivity by a random factor, as detailed in Section 4.2.4 for details. Gabaix (2009) identifies this stochastic process featuring a multiplicative component as an adequate generating mechanism for the empirically observed power-law distribution of firm size.

Given the resulting heterogeneous and dynamic levels of productivity, market shares are allocated accordingly. Our proposed allocation mechanism makes use of a biological metaphor, the Darwinian ‘survival of the fittest’ principle, now in the form of a ‘replicator dynamics’ approach (Fisher 1930). The fittest or most productive firms grow to dominate the market, while less productive firms fall victim to competition and are driven out (Cantner 2017). In our formal description, we remain agnostic about the precise nature of the mechanism translating productivity increases into growing market shares to allow for a reasonably general application. These translation mechanisms by which higher market shares might accrue due to enhanced productivity include: increased product quality for given unit costs; decreasing unit costs for products of equal quality; freed up funds for increased marketing spending; or any other mechanism.

---

<sup>1</sup>The focus on the upper tail is partially motivated by the fact that about a third of variations in US GDP growth can be explained by the idiosyncratic destinies of the 100 largest firms (Gabaix 2011).

<sup>2</sup>Here, we only describe the basic notion of the benchmark model and our extension. For details and all equations, see Section 4.3.

More specifically, in our representation the market share of a firm grows or shrinks according to how its productivity compares with the weighted average productivity of all firms in the model; thereafter, firms whose shares have fallen below a threshold leave the market and are replaced by new entrants. This constitutes a selective replicator dynamics process for which Cantner and Krüger (2008) as well as Cantner et al. (2012) present empirical evidence.

Dosi et al. (2017b) use these replicator dynamics in their model: Initially, all firms have equal market shares and productivity levels. At each time step, firms increase their productivity by a random factor, following which they gain or lose market share depending on how their own productivity compares with global average productivity. Firms whose share falls below a threshold value are replaced by new entrants. These new entrants start with the market share with which firms were initialised (and shares of all incumbents are adjusted so that the aggregate market size remains constant), but have their productivity level set to the current weighted global average productivity in the model. We carefully extend this model by adding a network layer to capture actual competitive interactions between agents. There are various ABM that emphasise the role of locality in economic interaction (Tesfatsion 2017). Our careful extension allows to gain more specific insights on the impact of localised competition without losing track of the core mechanisms. Hence, we validate the model by showing that the specific case of a complete network, which resembles the model by Dosi et al. (2017b), also displays a similar behaviour and yields the same stylised facts.

#### **4.2.4 Productivity Gain Through Stochastic Learning**

Our model features two channels of learning. Firstly, incumbent firms increase their productivity periodically. Secondly, the initial productivity of entrant firms depends on the localised market that they enter and hence they learn from their link neighbours.

The periodical learning describes the efforts of each firm to improve its productivity continually. In economics, the general concept of learning as a belief update justified by self-collected or socially acquired evidence (Zollman 2009) is often understood as a rational endeavour that agents explicitly control, as Evans and Honkapohja (2013) point out in their overview. Moreover, approaches such as Bray and Savin (1986) or Milani (2007) reveal a close connection between learning and rational expectations. While such detailed understanding of learning is appropriate when investigating a learning process itself, a macroscopic approach seems sufficient for a study of industrial dynamics, where the outcome of learning contributes to one of many mechanisms. This macroscopic approach focuses on the productivity gain that any learning activities of firms yield. Thus, the model abstracts from details of the learning process and does not distinguish where (e.g. product improvements, production efficiency, marketing) or how (e.g. new inven-

tions, imitation of others, deliberate management choices) the productivity gain takes place. When abstracting from subject-specific features of learning, one can treat success as being randomly distributed among individual learners and hence understand learning as an increase of productivity by a random factor. This stochastic learning seems to be an appropriate way to capture actual outcomes, as empirical findings are approximately represented (Luttmer 2007). Moreover, replicator dynamics are also consistent with an understanding of learning as imitating more successful others' behaviour. Schlag (1998) demonstrates this analytically by showing that the aggregate population behaviour follows a replicator dynamics whenever agents choose the individually most successful learning rule.

Accordingly, we follow Dosi et al. (2017b) and deliberately keep the learning process purely stochastic - do not explicitly include rational expectations or adaption to other firms - in order to focus on the network structure effects. With a purely stochastic process, we circumvent the problem that the precise form and effect of an innovation is per definition unpredictable and thus resort to much more modest statistical assumptions about the average rate of technological progress (Arrow 1991).

Besides the periodical stochastic learning of incumbent firms, the network layer and namely the localised market entry that it implies constitutes a second implicit mechanism of learning, which depends on asymmetric innovation. ABM studies that employ multiple or asymmetric learning processes in other contexts reveal unexpected system behaviour and have high explanatory power. For example, Klein and Marx (2018) and Klein et al. (2019) show in a model how asymmetric learning and information cascades shape individual estimates of how likely political revolution is. Asymmetric learning also plays a role in iterated games, as Macy and Flache (2002) show. Mayerhoffer (2018) runs a variant of the Hegselmann-Krause bounded confidence model (Hegselmann and Krause 2002) parallel to a network-based opinion update procedure and finds that the coexistence of both learning mechanisms can explain group-specific attitudes towards queerness among adolescents. In their model of Humean moral theory, Will and Hegselmann (2014) also employ explicit and implicit asymmetric learning in parallel. Models of learning and knowledge diffusion in networks also find application in business science, where they can provide explanations for competitive advantage, as Greve (2009) shows for shipbuilders and shipping companies. In these structures, Skilton and Bernardes (2015) find that successful market entry empirically depends on the network layout.

## 4.3 Model Description

This section provides a content-oriented presentation; for technical details see the commented model, which is appended electronically/publicly available at GitHub<sup>3</sup> and the description following the accompanying ODD protocol.<sup>4</sup>

Parameter	Variable	(Initial) Values	Equation
Population size	$N$	150	<i>constant</i>
Linking probability	$p$	0.01 – 1 (in steps of 0.01)	<i>constant</i>
Local market competitors	$K$	dependent on $p$	-
Firm Productivity	$a$	1	(4.1), at entry: (4.6)
Global market share	$s$	$1/N$	(4.2)
Localised market power	$l$	$1/K$	(4.4)

Table 4.1: Parameters, initial values and calculation. The interpretation of variables is given in Section 4.3.1.

### 4.3.1 Model Properties and Initialisation

The model observes a population of 150 firms that constitute an economy. We adopt this intuitively low number from Dosi et al. (2017b), but sensitivity analyses showed that our results also hold for larger populations. In this economy, firms try to maximise their sales revenue by improving their productivity through learning. However, whilst a firm does act rationally, this is only in a bounded manner due to its possession of imperfect information and environmental complexity. Hence, it does not adapt to the behaviour of other firms or form expectations. Undirected links connect some firms, but the firms themselves have no perception of their links.

Links between firms do not mean that they cooperate in research and development or in production; on the contrary, each link represents a direct competitive relationship between two firms in their selling of products that are (almost) perfect substitutes for each other. Pellegrino (2019) recently used the same methodology in a general equilibrium setting to identify aggregate trends and welfare costs of market power in the US. Thus, the model adds a level of locality to competition by linking firms. Clusters of densely linked firms represent an industry with aggravated competition. With this modelling approach, we combine two concepts of market structure that have enjoyed great success in the macroeconomic literature: Chamberlinian monopolistic competition (Chamberlin 1949; Robinson 1969) and the concept of a product space first introduced by Hidalgo and

<sup>3</sup>The computer simulation was implemented in Netlogo (Wilensky 1999).

<sup>4</sup>We describe the model according to the protocol guidelines by Grimm et al. (2010) and the extension proposed by Müller et al. (2013).

Hausman (Hidalgo and Hausmann 2009). From monopolistic competition, we take the notion that market power in differentiated, localised product markets is consistent with strong global competition, as indicated in the Chamberlinian concept by zero long-run profits. From Hidalgo and Hausmann (2009), we take the idea that similarity of products can be formalised by a network structure in a product space, where a network linkage indicates similarity.

In this context, to formalise local monopolistic competition within the product space, a Erdős-Rényi (ER) random network structure (Erdős and Rényi 1960) seems most appropriate because it does not call for an assumption that firms deliberately form competitive links. To generate the random network at initialisation, a link appears between any pair of firms with probability  $p$ . This probability is global, exogenously set and constant.  $p = 1$  represents the model by Dosi et al. (2017b). One major advantage of ER networks is the myriad of analytical results pertaining to the network structure for different  $p$ . We make use of two results in particular. Firstly, the degree distribution of the network is asymptotically Poisson with thin tails (Newman 2005). This result provides assurance that our findings on the fat-tailed nature of growth rates and the size distribution are no artefacts of the network structure we impose, but rather a genuine emergent feature of interaction within the model.<sup>5</sup> Secondly, there almost surely exists a single giant component for the range of network probabilities between 1% and 100% as the parameter range we consider. All other components have, almost surely, size of order  $O(\log(N))$ , where  $N$  is the number of firms. Hence, we can also examine the relevance of highly heterogeneous competitive environments, where the intensity of competition depends on whether or not a firm is connected to the single giant component or not (Erdős and Rényi 1960).

Besides its position in the network, each firm possesses three attributes that may vary over time. Firstly, there is the global market share  $s$ , which one can understand as sales revenue generated by each firm; it measures a firm's level of success and ultimately determines its survival. Initial shares of firms are equal:  $\forall i(s_i(t_0)) = 1/N$ . Secondly, local competition represented by the network structure means that firms also possess a localised market power  $l$ , which measures how productive a firm is in comparison with its immediate competitors that are the link-neighbours in our model. There is no immediate relationship between localised market power and sales revenue; high localised market power does not necessarily imply a high global share; for example, firms with great power in a small or unproductive industry may be small at the level of the whole market. Initial localised market powers are calculated following the same logic as global shares  $\forall i(l_i(t_0)) = 1/|K_i|$ , where  $K_i$  is the set containing all link-neighbours of  $i$  and  $i$  itself. Thirdly, productivity or level of competitiveness  $a$  of a firm indicates how well this firm is equipped for selling

---

<sup>5</sup>Note that even though empirical degree distributions are often fat-tailed warranting an explanation on their own (Johnson et al. 2014), there is no reason to impose a specific network structure a priori, as to the best of our knowledge, there exists no evidence on topologies of competition networks. We thus opt for as minimal assumptions as possible, leaving us with ER-type networks.

its products. It includes a firm's technological and business knowledge along with its skill base, but it could also be shaped by a specific demand for products that the firm offers. This attribute improves through learning and in turn, the productivity of a firm impacts the global share and localised market power of the firm itself as well as of other firms. However, for the purposes of our model, firms start with an equal level of competitiveness:  $\forall i(a_i(t_0)) = 1$ .

### 4.3.2 Events During the Simulation

The simulation proceeds in discrete time steps, within each of which the following processes take place in sequential order:

1. Learning: Firms (potentially) increase their productivity.
2. Assessment of global shares and localised market powers.
3. Entry and exit: Firms with low global market shares leave the market and new entrants replace them

**Learning** In the model, learning incorporates all processes that improve a firm's level of competitiveness. This includes an intentional quest for innovations in product design, efficiency of production, and supply-chain management as well as marketing. At the same time, according to this concept, a firm can also 'learn' if customers grow more interested in its products independent of deliberate actions by the firm (e.g., products become popular due to some trends set by third parties). We subsume this variety of aspects under a firm-specific and idiosyncratic learning mechanism that we take from Dosi et al. (2017b), while the general concept of this learning dates back to earlier work by Dosi et al. (1995) who propose this as a baseline condition in their model. For this multiplicative stochastic process, each firm  $i$  determines its productivity  $a_i$  as follows:

$$a_i(t) = a_i(t-1)(1 + \pi_i(t)) \quad (4.1)$$

Where  $\pi_i(t)$  describes a firm's learning parameter and is drawn from a rescaled beta distribution with  $\alpha = 1.0$ ,  $\beta = 5.0$ ,  $\beta_{min} = 0.0$ ,  $\beta_{max} = 0.3$  and an upper notional limit  $\mu_{max} = 0.2$ . This notional limit ensures that the maximum productivity growth rate is indeed 0.2, as we fix all drawn values higher than that for the notional limit. Firms do not experience negative learning because their productivity is measured in absolute terms rather than being compared with that of other firms at this stage. Learning in this sense does not entail failures which would imply negative productivity gains. This deliberate modelling choice excludes planning mistakes on the part of individual firm management, and isolates effects generated by the interplay between stochastic learning and market



selection. Furthermore, absolute productivity growth depends on the firm's previous productivity level, meaning that the expected productivity gain grows proportionally to its past productivity. However, learning is independent of firm size, and hence, there is no (direct) amplifier that would reward larger firms with a potential for higher rates of productivity gain.

**Assessment of Global Shares** A replicator dynamics formulation reproducing the one used by Dosi et al. (2017b) determines the global market share:

$$s_i(t) = s_i(t-1) \left(1 + \frac{a_i(t) - \bar{a}(t)}{\bar{a}(t)}\right) \quad (4.2)$$

For any market containing at least two firms, it holds that  $0 < s_i(t) < 1$ . The global parameter  $\bar{a}$  is the weighted average productivity of all firms  $N$  in the global market:

$$\bar{a}(t) = \sum_j a_j(t) s_j(t-1) \quad (4.3)$$

A firm's global share depends not only on its own productivity level, but also the weighted productivity levels of all other firms in the market. The weighting ensures that larger rather than smaller firms shape the market more distinctively. The sales revenue of a firm grows (or shrinks) according to how much the productivity level of this firm exceeds (or undercuts) the weighted average productivity level.

**Assessment of localised market power** The calculation of localised market power  $l_i$  is similar to that of global market share:

$$l_i(t) = l_i(t-1) \left(1 + \frac{a_i(t) - \bar{a}_i^L(t)}{\bar{a}_i^L(t)}\right) \quad (4.4)$$

However, agents now compare their productivity level only to those of their link neighbours and the firm itself, the set  $K$ :

$$\bar{a}_i^L(t) = \sum_{j \in K_i} a_j(t) l_j(t-1) \quad (4.5)$$

Consequently, a completely unconnected firm (i.e., one with a local monopoly on all its products) has a localised market power of 1, while the value for each other firm may be above or below its global share.

**Entry and exit** When the global share of a firm drops below the threshold of 0.001, it leaves the market. Using the global share rather than the localised market power here makes sense because a firm becomes unprofitable or goes bankrupt if its sales revenue is

too low. This may happen even to firms that possess high power in an industry which ultimately proves to be too small and unsustainable, while conversely firms that are small players in a large industry may create a high revenue. Firms leaving the market disappear from the model and also destroy all their links, which improves the local position of their old link-neighbours (i.e., former direct competitors).

Each departing firm is immediately replaced by a new entrant. This entrant links with incumbents and other new entrants with the same probability  $p$  used for initial network generation. The global share of entrants is  $1/N$ , and their localised market power is  $1/|K_i|$  (with  $K_i$  again being the set of all link-neighbours and the entrant  $i$  itself). However, entrants do not start with a low productivity value of 1, but instead acquire the specific productivity level of the industry they enter (or of the whole market as a fallback should the entrant have no links) altered by the common learning parameter:

$$a_i(t_{\text{entry}}) = \begin{cases} \overline{a}_i^L(t_{\text{entry}}) + \pi_i(t_{\text{entry}}) & \text{if } K_i \setminus i \neq \emptyset \\ \overline{a}(t_{\text{entry}}) + \pi_i(t_{\text{entry}}) & \text{if } K_i \setminus i = \emptyset \end{cases} \quad (4.6)$$

Generally, entrants benefit from past technological and management innovations as well as local market conditions that shape competitiveness in the industries that they enter. Locally stronger incumbents in the industry play a more central role here. The productivity level of link-neighbours weighted by their localised market power reflects this notion. Here, localised market power is used as the weighting instead of global share, because the local importance of an incumbent firm matters for entrants and not absolute sales revenue. This again reflects the boundedly rational nature of our model, where entrants learn indirectly from the experience of local incumbents, particularly in regard to their tacit product domain knowledge (Glauber et al. 2015). Some entrants are uniquely innovative, meaning that they increase their own productivity levels beyond those of their environments.

Because entrants' global shares are higher than those of market leavers, the sum of all global shares now exceeds 1. To correct for this, the share of each firm is reduced proportionally (divided by the sum of all shares). Likewise, the entry/exit process altered the network structure and thus each firm divides its localised market power by the sum of all localised market powers in  $K$ . These adjustments normalise the corresponding values and ensure comparability over time.

## 4.4 Simulation Results

To ensure comparability, the results presented in this section generally follow the parametrisation in Dosi et al. (2017b): In particular, we opt for 50 Monte Carlo iterations of simulation runs with time-series length  $T = 200$  and for  $N = 150$  firms each.

Given that the trajectories of all aggregate statistics for all linkage probabilities converge to a stationary state very quickly, we present our estimations for  $t = 200$  only, but for the pooled Monte Carlo simulation runs. For  $p = 1$ , which in its model assumptions fully corresponds to the model set-up in Dosi et al. (2017b), we derive results that are also in full qualitative agreement with the results obtained by Dosi et al. (2017b), indicating that we are indeed generalising their case.<sup>6</sup> For all  $p$ , we find that the productivity distribution is very heterogeneous with fat tails and hence consistent with Dosi et al. (2017b) and the empirical evidence cited therein. The negative relationship between size and variance in growth rates, though, holds only for  $p = 1$  and its neighbourhood, strengthening the case we are building in this section to confirm that the fully connected network is indeed the empirically relevant benchmark.<sup>7</sup> In contrast to non-parametric descriptive statistics, we analyse our three main results by fitting parametric distributions to the data that allow us to examine the generating mechanisms pertaining to the variable in question.

#### 4.4.1 Size Distribution

We find that the upper tail of the firm size distribution (measured by market shares) is for all  $p$  characterised by a power-law distribution. The distribution cannot be statistically distinguished by standard non-parametric tests such as the Cramér-von-Mises test and the Kolmogorov-Smirnov test from a continuous power-law distribution (Anderson and Darling 1952; Anderson 1962; Smirnov 1948). We report the corresponding test statistics and p-values in Table 4.2 below.

Furthermore, for all  $p$ , the power-law regime spans approximately two orders of magnitude and hence meets the common minimum requirement for a power-law to be present (Stumpf and Porter 2012). This is equivalent to saying that the discrete distribution of shares can be approximated by this Probability Density Function (PDF) for the continuous analogue of this power-law distribution as

$$p(s) = Cs^{-\alpha}, \text{ for } s \geq s_{min}, \quad (4.7)$$

where:  $s_{min}$  denotes the minimum share from which on the power-law applies;  $\alpha$  denotes the characteristic exponent of the power-law distribution, and  $C$  is a normalising constant letting the probability density integrate to 1. Notice that  $\alpha \geq 1$  is an inverse measure of concentration, where a lower  $\alpha$  indicates a higher degree of inequality. Figure 4.1 shows the Complementary Cumulative Distribution Function (CCDF) of the upper tail firm size distribution for three different linkage probabilities. The minimum was determined by the standard procedure in this field first outlined by Clauset et al. (2009), essentially by:

---

<sup>6</sup>We are strongly indebted to Giovanni Dosi and Marcelo Pereira for their helpful comments and sharing their code with us to establish this comparability in all subtle details.

<sup>7</sup>The results on the variance-size scaling, as well as the productivity distribution, are available upon request.

<b>p</b>	<b>Test Statistic KS</b>	<b>p Value KS</b>	<b>Test Statistic CvM</b>	<b>p Value CvM</b>
0	0.0496	0.327	0.265	0.171
0.1	0.0165	0.785	0.062	0.802
0.2	0.0225	0.347	0.140	0.421
0.3	0.0218	0.356	0.161	0.357
0.4	0.0124	0.930	0.0428	0.918
0.5	0.0210	0.569	0.0939	0.615
0.6	0.0178	0.935	0.0501	0.876
0.7	0.0196	0.452	0.194	0.279
0.8	0.0149	0.860	0.066	0.776
0.9	0.0211	0.515	0.152	0.383
1	0.0565	0.143	0.320	0.119

Table 4.2: Summary of test statistics and estimated p values for a fitted power-law distribution to the simulated sales shares with  $p$  in increments of 0.1 for the whole parameter space. “CvM” refers to the Cramér-von-Mises test, and “KS” to the Kolmogorov-Smirnov test. For both tests and all reported linkage probabilities  $p$ , the hypothesis of a power-law cannot be rejected at the usual significance level of 5%.

fitting a reverse-order-statistic to a power-law with increasing sample sizes; obtaining the Kolmogorov-Smirnov test statistic; and choosing the  $s_{min}$  that minimises it. This method has been shown to outperform other methods robustly, such as minimising the Bayesian Information Criterion (Clauset et al. 2009). Indeed, as the non-parametric tests also suggest, all three distributions display an approximately linear behaviour on a double-logarithmic scale (Newman 2005). The slope for each  $p$  is approximately  $\alpha$ , where a lower  $\alpha$  indicates that the CCDF does decay more slowly, indicating a higher concentration with a higher frequency of large shares.

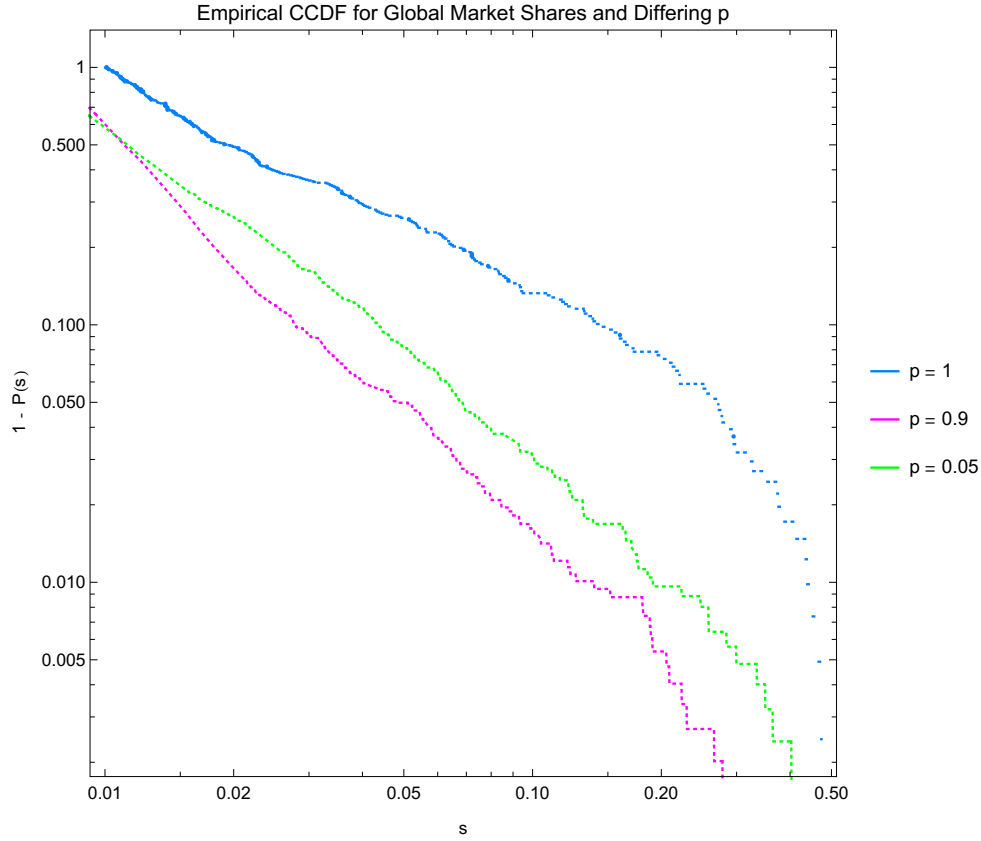


Figure 4.1: Empirical Complementary Cumulative Distribution Function of Firm Sizes for  $p = 0.05$ ,  $0.9$  and  $1$ .

Given the set-up of our model, this is not surprising as it essentially comprises a stochastically multiplicative process with an entry-exit mechanism that has been shown to be the most promising candidate for generating power-laws (Gabaix 2009). Hence, for our set-up the path-dependent stochastically multiplicative process seems to remain the most critical feature of the model, irrespective of the underlying network structure. Moreover, regardless of the underlying mode of local competition, we want to highlight that this extremely heterogeneous power-law distribution implies a situation that is far from the perfect competition usually assumed as a benchmark for general equilibrium models.

The functional form for the upper tail of empirical firm size distributions is thus seemingly broadly consistent with all connectivity patterns for the underlying localised network. However, the empirical consensus that the upper tail of firm sizes is characterised by Zipf's law with an estimated  $\hat{\alpha}$  not statistically different from 1 constrains the permissible  $p$  to a much more narrow range. In Figure 4.2, we show the behaviour of estimated  $\hat{\alpha}$  for the whole range of  $p$  in our model using 1% increments and with error bands corresponding to two sample standard deviations upwards and downwards, implying that the plotted intervals span the true  $\alpha$  with 95% confidence.<sup>8</sup>

<sup>8</sup>The estimation of  $\hat{\alpha}$  was carried out by using the associated Maximum Likelihood Estimator (MLE) or Hill estimator that has been shown to be less biased compared to OLS methods or fitting a linear

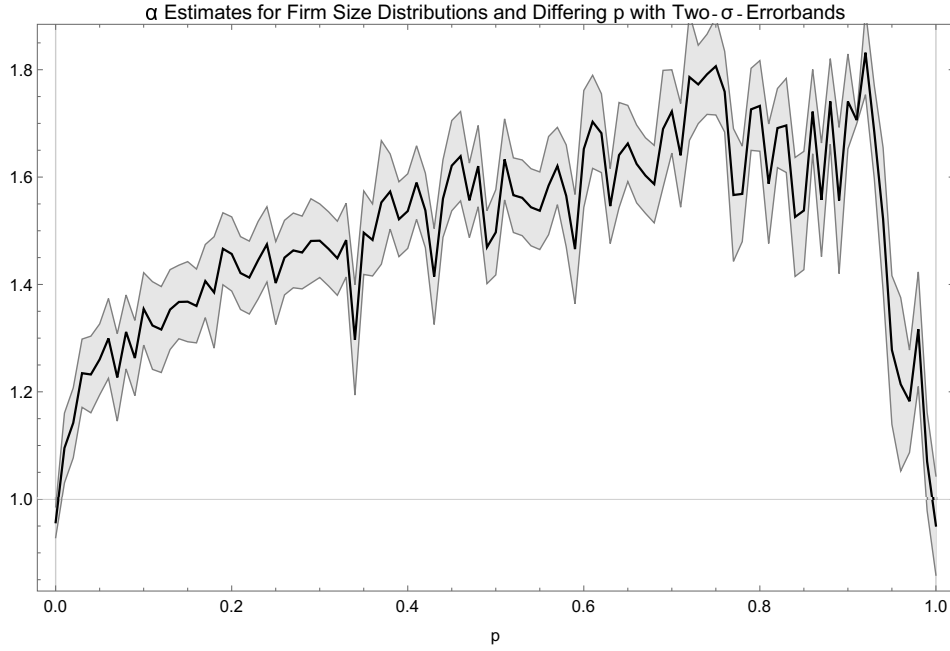


Figure 4.2: Estimated  $\hat{\alpha}$  for all  $p$  in increments of 1%. Lines between estimates are visual aids only.

Two features are striking in the plot: Firstly, Zipf's law is consistent only with the two knife-edge scenarios of an extremely sparse network in the (narrow) neighbourhood of  $p = 0$  and the other extreme of a very dense network in the (narrow) neighbourhood of  $p = 1$ . This, in turn, implies that the empirical evidence constrains us to these two extremes. Secondly, contrary to economic intuition built within general equilibrium models, measured concentration is maximal - Zipf - for the highest degree of local competition and lowest for a mild (local) oligopoly around  $p = 0.9$ .

#### 4.4.2 Growth Rates of Market Shares

Another focal point of the industrial dynamics literature is the presence of fat-tailed growth rate distributions in sales. In more colloquial terms, this implies that jumps in firms' market shares are relatively more frequent than one would expect from a Gaussian distribution. Note that the presence of non-Gaussian growth rate distributions alone indicates that the growth process is not independent in time. According to the Central Limit Theorem, this would induce Gaussian growth rates. Of course, stochastically multiplicative growth processes like ours responsible for emergence of the power-law in levels are actually path-dependent and thus violate independence. Indeed, Dosi et al. (2018) produce robust findings which support fat-tailed growth rates for the baseline specification

function onto the power-law on a double-logarithmic scale. Cf. also Goldstein et al. (2004) for a more rigorous analysis of different graphical methods and their respective shortcomings compared to an MLE. The standard errors were obtained exploiting the asymptotic Gaussianity of the Hill estimator (De Haan and Resnick 1998).

we use together with a vast range of different specifications and parameter constellations. Their baseline model corresponds to our  $p = 1$  parametrisation. We also find fat-tailed growth rate distributions for  $p$  different than 1, as can be seen for  $p = 0.05$  and 0.9 in Figure 4.3. However, at least for the not fully connected network, the fat-tailed nature of growth rates is primarily due to extreme losses, rather than frequent extreme growth events which are at odds with the presence of superstar firms. We want to highlight also that this fat-tailed nature is a different concept from mere ‘dispersion’. While dispersion does indeed seem to decline with  $p$ , being fat-tailed refers to the frequency of extreme events *relative* to the frequency of events closer to the expected growth rates, for which inference by visual inspection is a much harder task. Within Figure 4.3, frequent extreme growth events are present only for the fully connected network. This finding, though, might merely be an artefact of the three network connectivities under consideration and thus will not hold for the whole parameter space. We need to explore the full parameter space to identify possible switching behaviour concerning the source of fat tails in the simulated growth rate distributions.

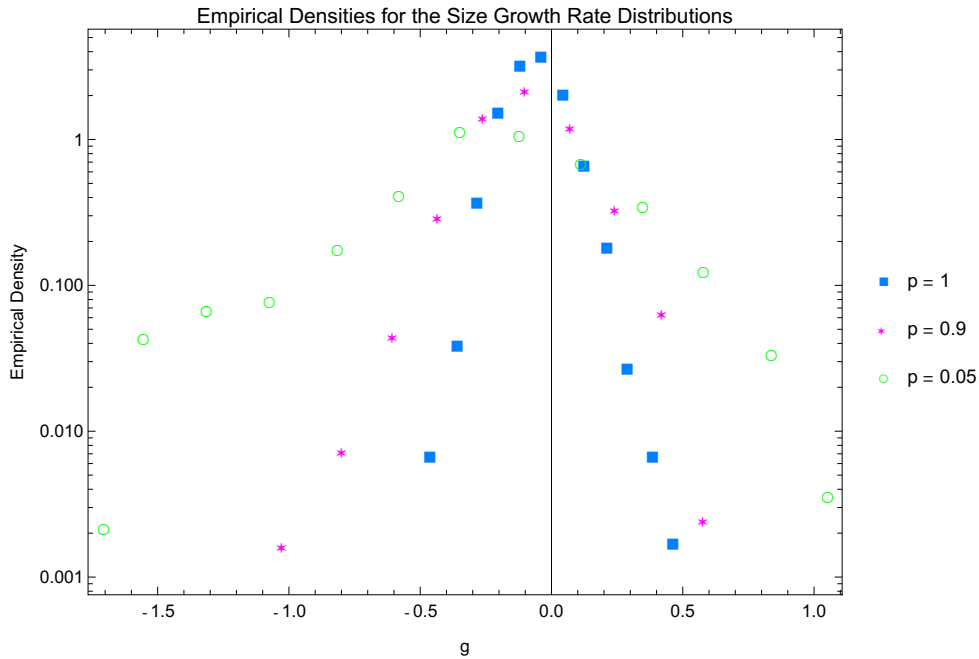


Figure 4.3: Empirical Density of Firm Size Growth Rates  $g$  on a semi-logarithmic scale for  $p = 0.05$ , 0.9 and 1.

A standard procedure for identifying fat tails and quantifying the degree of ‘fat-tailedness’ in growth rate distributions is to fit a Subbotin distribution (Subbotin 1923) to the data and take its shape parameter  $b$  as a measure of heavy tails’ strength. The Subbotin density includes the Gaussian for  $b = 2$ , the Laplacian for  $b = 1$ , the Dirac-Delta for  $b \rightarrow 0$  (from above), and the uniform distribution for  $b \rightarrow \infty$  as special cases. Consequently, we define fat-tailed behaviour for all  $b \geq 0$  significantly smaller than 2 for the Gaussian case. As the

contemporary relevance of superstar firms is central to our concerns within this study, we are primarily interested in extreme growth events as opposed to extreme losses. We opt for an asymmetric variant of the Subbotin distribution, introduced by Bottazzi (2014), to distinguish extreme growth events from extreme losses. The PDF is given by:

$$p(g; a_r, a_l, b_l, b_r, m) = C^{-1} \text{Exp} \left( -\frac{1}{b_l} \left| \frac{g-m}{a_l} \right|^{b_l} \Theta(m-g) + \frac{1}{b_r} \left| \frac{g-m}{a_r} \right|^{b_r} \Theta(g-m) \right), \quad (4.8)$$

where:  $\Theta(\cdot)$  denotes the Heaviside theta function;  $m$  is a centrality parameter;  $a_l$  and  $a_r$  are the scale parameters of the left and right tails, respectively; whilst  $b_l$  and  $b_r$  are shape parameters for both tails with the analogous interpretation as in the symmetric case. In the language of this distributional analysis, ‘superstar-like’ behaviour is obtained for relatively frequent extreme growth events, that is, a fat right tail of the growth rate distribution with  $b_r$  significantly lower than 2. We estimate both parameters for the growth rate distributions of market shares by MLE for each  $p$  in 1% increments. The corresponding standard errors are obtained by utilising the Fisher information (Ruppert 2014).<sup>9</sup> Figure 4.4 shows the values of  $b_l$  and  $b_r$  as a function of  $p$  with  $p$  increasing in 1% increments.

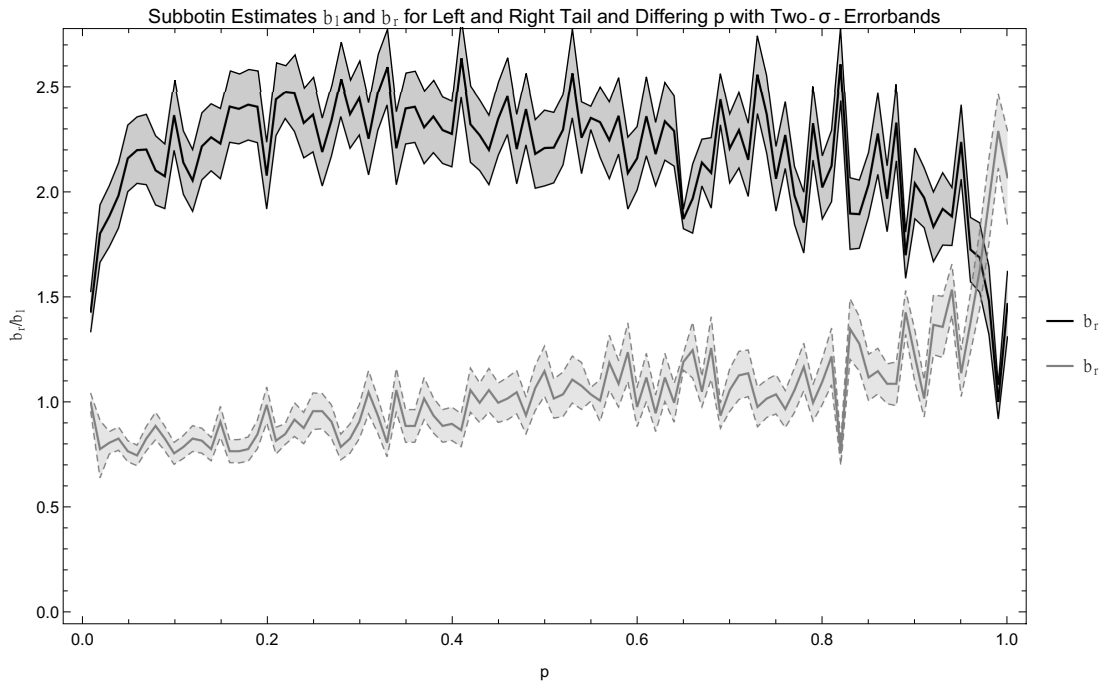


Figure 4.4: Estimated  $\hat{b}_l$  and  $\hat{b}_r$  for all  $p$  in increments of 1%. Black estimates for  $b_l$ , gray estimates for  $b_r$ . Lines between estimates are visual aids only.

<sup>9</sup>In particular, we employ the freeware *Subbotools 1.3.0* specifically designed for the estimation of different flavours of the Subbotin distribution (Bottazzi 2014), which delivered by far the most efficient parameter estimates for different samples of data we simulated.



The figure highlights two distinct regimes with respect to the growth rate distributions. Taken as a whole, all growth rate distributions seem to be fat-tailed in agreement with empirical studies. However, the source of this fat-tailed behaviour differs between regimes. While for the broadest range of  $p$  between 0 and about 0.93, thus between a completely sparse and a very dense network, relatively frequent *extreme losses* are responsible for the fat tails, the situation changes dramatically in the neighbourhood of a fully connected network, where relatively frequent *extreme growth* dominates. Superstar-like behaviour is thus consistent only with extremely dense networks implied by a  $p$  close to 1.

### 4.4.3 Age

Finally, for age, our model can mimic the empirically observed exponential distribution in age levels for all  $p$ . This can be seen in Figure 4.5, where the three age distributions being considered display approximately linear behaviour on a semi-logarithmic scale, consistent with an exponential functional form.<sup>10</sup> This emergent exponential stationary age distribution coupled with stable population levels has, in itself, an important implication: The exit probability or insolvency rate is common and constant for all firms (Artzrouni 1985). Thus, every firm irrespective of its age has the same probability of becoming insolvent in any period and, consequently, has the same expected age. The relatively stable distributions of size and growth rates within time at the meso-level are consistent with a very dynamical economic system underlying these regularities and high rates of ‘churning’ in the composition of firms, where even local and global market leaders face the same certain prospect of insolvency at some point.

---

<sup>10</sup>The exponential might appear to not fit the age distribution too well for  $p = 0.9$ . However, this impression is mainly an artefact of the semi-logarithmic scale and pertains only to the largest 0.1% of values. For the remaining 99.9%, the fit is extremely good, leaving us confident that the exponential is a reasonable choice here.

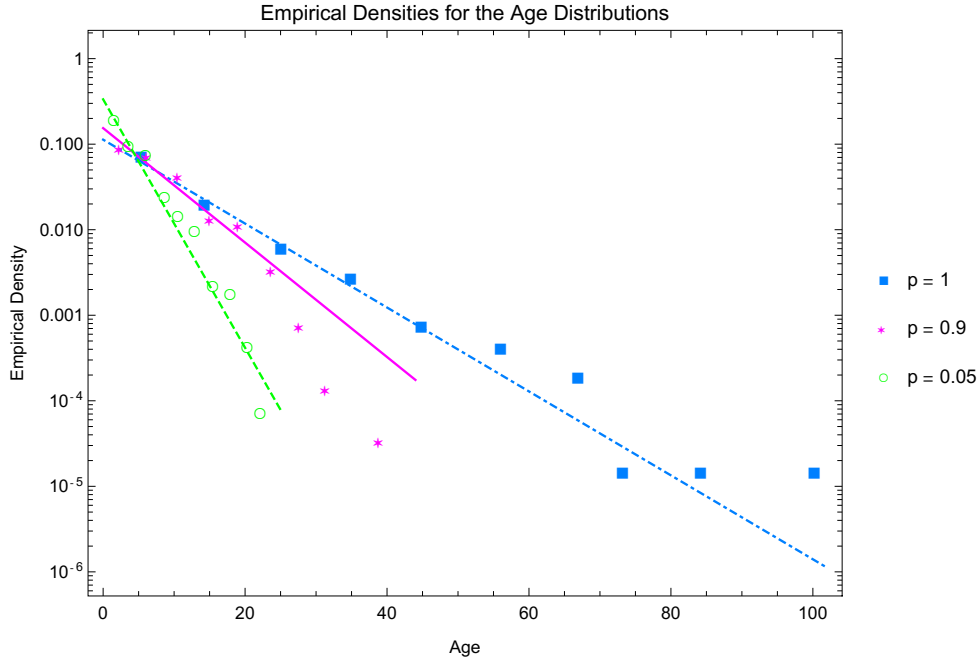


Figure 4.5: Empirical Density of Firm Age on a semi-logarithmic scale for  $p = 0.05, 0.9$  and 1 with Exponential Fits obtained by MLE.

While the functional form of the emergent age distribution is constant for different  $p$ , its estimated parameter  $\hat{\lambda}$  as the insolvency rate changes with  $p$ . As can be seen, the firms' life expectancies vary widely with  $p$ . For  $p = 0.05$ , the firms cluster around a very young age, while exhibiting much wider dispersion and higher expected age for  $p = 1$ . Life expectancies thus seem to increase in the network connectivity, but also grow more heterogeneous.

When examining the whole parameter space of  $p$ , the insolvency rate falls monotonically with  $p$  as we show in Figure 4.6.<sup>11</sup>

Thus, while all firms irrespective of their age face the same estimated insolvency probability  $\hat{\lambda}$  per regime or per  $p$ , this insolvency probability differs widely between the regimes. Given that expected age is just the inverse of  $\hat{\lambda}$ , this implies that with a higher  $p$ , firms tend to stay in the market for a much longer time, and there is less 'churning' between periods.

#### 4.4.4 Generating Mechanisms

Dosi et al. (2017b) explain their model outcomes based on the idiosyncratic learning process and replicator dynamics; this explanation also straightforwardly applies for the complete network in our model. However, the tails of the global share distribution, the

<sup>11</sup>The estimated  $\hat{\lambda}$  was estimated through MLE. The standard errors were obtained by utilising the fact that  $\hat{\lambda}$  is just the inverse of the sample mean and that the associated sample standard deviation is therefore  $(\hat{\lambda}\sqrt{N})^{-1}$  (Lehmann and Casella 2006).

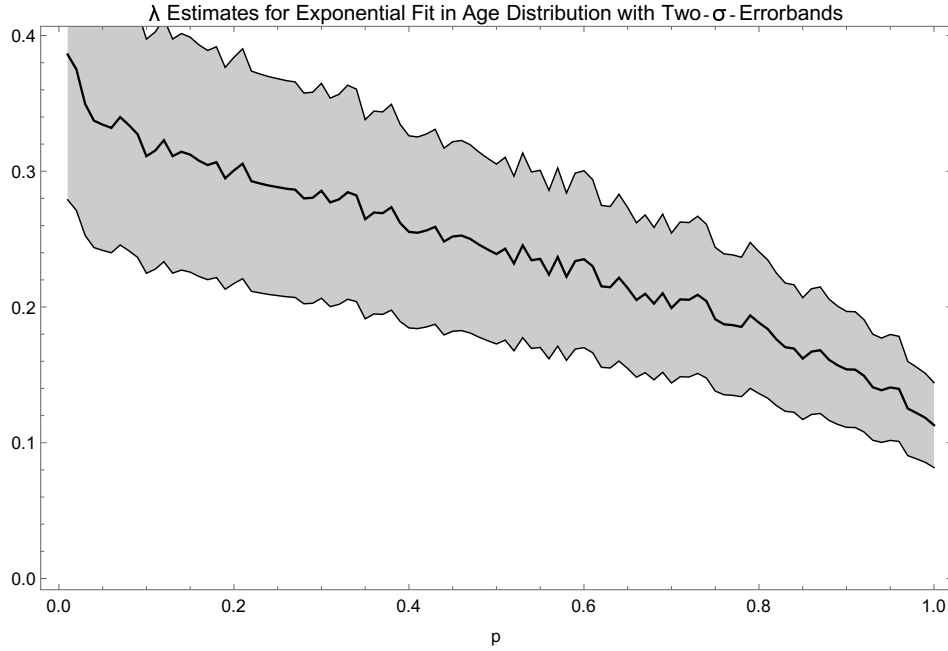


Figure 4.6: Estimated  $\hat{\lambda}$  for all  $p$  in increments of 1%. Lines between estimates are visual aids only.

growth distribution and market exit probability of firms react in a highly elastic way towards changes in the network topology. Thus, these results suggest the presence of a second model mechanism that depends on network density and the implied distribution of localised market power. Because the learning of incumbents and assessment of their global shares work irrespective of network layout, the success of entrants remains the sole candidate for such a driving mechanism.

Since all entrants have identical initial market shares, the individual success of each entrant depends largely on its initial level of competitiveness. To gain a high level of initial competitiveness, an entrant must connect to as many highly productive incumbents with high localised market power as possible (i.e., join a thriving industry). Such connections become less likely for smaller link probabilities  $p$ . Thus, in sparse networks, most entrants start with low productivity. Furthermore, since the assessment of global market share compares the productivity of the firm in question with the weighted global average productivity, these relatively unproductive entrants quickly lose market share in the first periods of their lives. That explains the fat left tail of the growth rate distribution for small  $p$ .

At the same time, those entrants that connect to highly productive and powerful incumbents have a comparative advantage, increase their sales quickly, and manage to catch up with even the most successful firms in the market. Hence, firms within the power-law tail exhibit more homogeneous sizes, the sparser the network is. The low maximum firm age and high probability of exit for sparse networks is a corollary of these two aspects:

Even successful firms are challenged, find themselves outperformed by productive younger competitors and finally leave the market, while most entrants do so after only a few simulation periods. Without explicit targeting, we are able to replicate ‘imprinting’ behaviour or the empirically well established phenomenon that founding conditions exhibit lasting effects on the entrants’ survival probabilities (Geroski et al. 2010).

For higher linking probabilities, the rate of entrants with a high initial productivity level grows, making the left tails of the growth distribution thinner. However, the most productive entrants are also hindered by the higher average productivity level and consequently have a harder time becoming superstars; thus, the inequality within the power-law tail decreases even more.

Furthermore, the importance of the birth productivity mechanism, which favours few entrants and lets many suffer, becomes weaker the denser the network is and hence the more similar localised market power as well as global market share become. If the entrants’ fate is no longer determined at birth, learning becomes more important. Thereby, the replicator dynamics of global share assessment means a fat right tail of the growth distribution and tails of the firm size distribution in accordance with Zipf’s law.

To summarise, two distinct mechanisms govern the productivity of firms and consequently their commercial success. The first is a process of learning that occurs within each period and is equally strong for all network layouts, but its effects depend on attendant productivity levels. The second mechanism is the allocation of initial productivity based on link-neighbours, which applies only once to each entrant at birth. The mode of operation and the strength of this second mechanism depends to a great extent on network density. For least dense networks, it dooms most of the entrants to a fast market exit while it is at the same time also subsidising a few of them in an extreme way, prolonging their accumulation of market shares. For denser networks, more firms share this subsidy and hence the most successful firms become more equal in terms of their size. Furthermore, it is noteworthy that the second mechanism takes precedence over the first for all but the densest networks, according to simulation outcomes. That is the case because birth productivity also implies a path-dependency: For an unequal birth productivity distribution of entrants, learning stabilises, and amplifies this inequality due to a higher productivity, also meaning a potential for a higher absolute gain through learning.

## 4.5 Discussion

We introduce a network-structure to the bare-bones model of a ‘winner takes most/all’ market proposed by Dosi et al. (2017b). This extension generates surprisingly rich dynamics and intriguing implications when deviating from the benchmark of a fully competitive localised market. In particular, we have been looking to highlight both the positive and

normative implications we draw from our modelling exercise and their practical relevance for economic regulation as well as management decisions in the case of a single firm.

Empirically, we find that the stylised facts of industrial dynamics, namely Zipf’s law in the firm-size distribution and fat-tailed, ‘superstar-like’ firm growth rates are consistent only with a situation very close to the benchmark of a fully connected network, meaning most intensive localised competition. All other network connectivities lead to significant deviations from the stylised facts in at least one regard. Hence, if we can accurately identify parts of the empirical mechanism – and there exists evidence that replicator dynamics play an essential role in empirical markets (Cantner and Krüger 2008; Cantner et al. 2012), our results will point to product markets that are relatively undifferentiated. Thus, market power comes from global rather than localised dynamics. These results are in stark contrast to our initial expectations of low concentration and high rates of ‘churning’ for relatively high degrees of localised competition. This indicates that anecdotal insights gained from analysing static frameworks of competition do not necessarily transfer well to situations where strong non-linearities and feedback mechanisms are present.

In our model, the coexistence and partial interaction of two learning mechanisms and replicator dynamics explain the results: (1) stochastic productivity improvements in each period for each incumbent firm constitute the first way of learning; (2) works indirectly at market entry because an entrant’s initial productivity depends on the weighted average productivity of the incumbents that it links to, meaning within its specific industry. The less densely connected a network, the fewer entrants form connections to highly productive incumbents; hence, their initial productivity is low, and consequently, their market shares decrease, which explains the fat left tail of growth rate distributions and lower average firm age. However, those entrants connected to highly productive incumbents thrive because their initial productivity is high in comparison to most incumbents. Thus, they can catch up with even the most successful incumbents and market concentration decreases. Methodologically, our model of networked competitive interaction can thus be thought as a complement to the foundational theoretical study by Cantner et al. (2019) who study collaboration in networks to explain especially the instability of early-lifecycle firms by lock-in effects within suboptimal value chains. Namely, our model suggests a mechanism that may be present in addition to ‘failures of selection’ (Cantner et al. 2019) and cause the high rates of churning and volatility in market shares of young firms already observed by Mazzucato (1998). We demonstrate that such instability can also emerge for functioning selection and industry-specific initial productivity, as long as markets are locally segregated or, equivalently, the competitive network exhibits rather low density. We aim to investigate the interplay of both the collaborative and competitive network channels in further research.

Since inequalities of initial productivity shrink with increasing network density and do not exist for the complete network, a lower level of competitiveness implies lower market con-

centration. Furthermore, for the complete network, other sources of inequality in initial productivity could replace the market entry learning mechanism: The absence of inequality in starting conditions leads to the most successful firms acquiring a greater market share. Consequently, one must accept the success of these superstars as an outcome if the aim is to create full equality of opportunity; otherwise, to avoid high market concentration by the most successful firms, one must deliberately create inequality of opportunity. In less abstract terms, our model suggests the common fear that active industrial policy creating unequally favourable starting conditions for specific firms and thus being anti-competitive to be at least partially misguided (cf. Sokol (2014) for a vocal proponent of this view): The relevant metric for consumers is perhaps *ex post* concentration in market shares, indicating that one can accept or even foster *ex ante* inequality in starting conditions to decrease such concentration after the fact. Active industrial policy enhancing the productivity of incumbents can even lead to positive productivity spill-overs, since entrants benefit from the average productivity of the market they enter. The more relevant trade-off within such markets appears to be between decreasing concentration in market power (lowering  $p$ ) or decreasing the amount of turnover in the market (increasing  $p$ ), with ‘turnover’ typically also implying (transient) increases in unemployment and the destruction of firm-specific capital and knowledge. In this way, our model can help to identify the relevant trade-offs for regulatory policy and contribute to a richer view apart from standard static efficiency considerations.

Besides these global findings, the model also suggests that there are localised cycles of productivity and firm size: If incumbents have acquired high localised market power and a high productivity level, new entrants joining the market segment and engaging in competitive interaction (i.e., linking to the productive incumbents) also start with a high productivity rate. Consequently, the industry in question becomes even more productive until it overheats, and incumbents are repressed from the market while the high productivity shifts to another (possibly new), related market segment. This effect is entirely in line with the empirical study, in which Schlichte et al. (2019) show that the timing of entry to highly specified submarkets between two technology waves is crucial for the success of new firms. Moreover, our model supports their finding that there is a first-follower advantage (in our model represented by successful entrants) as opposed to a first-mover advantage (moderately successful incumbents that are nonetheless outperformed by entrants) because of growing consumer acceptance of new technology (Davis 1989). These findings might be of particular interest to practitioners in Venture Capital and are consistent with their empirical emphasis on ‘deal selection’ compared to other phases of the investment process (Gompers et al. 2020). However, there exists no consensus on the correct selection strategy, with some trend-following venture capitalists selecting ‘hot sectors’ and other contrarian ones avoiding them (Gompers et al. 2020). In principle, our model points to the trend-following strategy to benefit from the high initial local

productivity of the relevant submarket. This is still no guarantee for success, though, as the submarket in question might be on the brink of overheating, also providing a rationale for the contrarian view. Venture capitalists, in our model, should thus pick sectors with high expected *growth* in contrast to present size in *levels* to avoid entering markets near the end of a technology wave.

The validity of our model depends to a great extent on the validity of the baseline model by Dosi et al. (2017b), which we assume to be given. However, since we re-implement the mechanisms from scratch and include the baseline model as a special case reproducing its findings, we can affirm the internal validity of the baseline model and our extension. With regard to external validity, we hope that including a network structure of localised competition can facilitate resemblance (Mäki 2009) between model and real-world economies. Our explanans can actually be true and the cause for the observed empirical fact. Hence, our proposed mechanism fulfils the minimum conditions for a good epistemically possible how-possibly explanation formulated by Grüne-Yanoff and Verreault-Julien (2021). Nevertheless, with the inclusion of localised market power, the nature of our model and thus, the implied mode of analysis remains highly stylised. Hence, the validity of the model is based on its “qualitative agreement with empirical macrostructures” (Fagiolo et al. 2019, p. 771), namely the replication of the stylised empirical facts that our model successfully attempts. Put differently, we develop a specific parallel reality (Sugden 2009) that features generating mechanisms for empirical findings in our reality and hence our results present a candidate explanation for the stylised empirical facts (Epstein 1999). Consequently, there may be different, more adequate, parallel realities featuring either these or even better mechanisms, despite to the best of our knowledge there being no existing models that fulfil these characteristics.

Alternative mechanisms firstly concern the network that we use. While we test for any network density, we limit ourselves to random link formation as we are not aware of empirical evidence for any specific network topology in our context. However, a non-random (e.g., preferential attachment or spatial-dependent) link formation may impact simulation results, especially for low network densities. Moreover, we distinctly interpret links as indicators for localised competition that only matters for a firm’s initial productivity level. One could further explicate such localised competition and track it over time. Alternative or additional layers of links could also represent cooperation between firms or their products being complements. Our model’s most apparent limitation concerns the baseline replicator dynamics equation, though, which implies that the emergent concentration is ‘good concentration’ (Covarrubias et al. 2020) and fully justifiable by productivity differences. Empirically, it is questionable if concentration indeed only reflects productivity (Covarrubias et al. 2020), with firms erecting artificial barriers to entry or acquiring competitors and discontinuing their innovative product lines in so-called ‘killer acquisitions’ (Cunningham et al. 2021) leading to ‘bad concentration’. Since it is at least conceivable

that a high concentration of the good type is preferable to lower bad concentration, the inclusion of strategic anticompetitive behaviour might alter the policy conclusions of our model and tilt them more towards antitrust measures, which might like in our baseline model induce high ex-post concentration purely based on productivity differences.

Furthermore, the stylised nature of our findings points to obvious extensions and avenues for further research. To analyse not only the consistency with stylised facts, but also for quantitative predictions and even policy experiments, the ABM community has recently developed certain new methods. These are aimed at bringing modelling closer to the data and calibrating parameters (Hassan et al. 2010), particularly by utilising the Method of Simulated Moments (Gourieroux and Monfort 1996) and related approaches (Bargigli et al. 2020). Given the partial equilibrium nature of our model, this would probably also necessitate allowing for a variable total number of firms over time by including mergers & acquisitions as well as consumer demand, a state sector and even financial markets for meaningful policy experiments. With the benchmark model by Dosi et al. (2017b), this was attempted by the K+S ABM (Dosi et al. 2010). An extension of this sort would enable us in future research to conduct policy experiments and quantify welfare effects for different market structures.

We believe our model to be a valuable contribution to the discussion on market structures and a key step towards a unifying explanation for both the microeconomic evidence on ‘superstar’ firms and the distributional findings in industrial dynamics. To the best of our knowledge, these two strands of literature have not as yet been linked. We find that for the replicator dynamics approach from industrial dynamics to be consistent with the existence of superstar firms, shown by microeconomic studies, there needs to be an almost perfect level of competition. These outcomes emerge because in each simulation period, the firms improve their productivity by idiosyncratic stochastic learning, while new entrants adopt the specific productivity level of their sub-market (i.e., the productivity of their immediate competitors weighted by localised market power). Thus, the model suggests that new firms are most successful if they join existing, highly productive submarkets with high growth potential. Accordingly, while accepting bounds on rationality, we can single out the strategy which a perfectly rational market entrant would pick when faced with a certain market structure. This not only fosters an understanding of market dynamics, but can also be applied to highlight the importance of market intelligence for the management in new firms and new technology markets.

## References

- Akerman, A., Leuven, E., and Mogstad, M. (2022). Information frictions, broadband internet and the relationship between distance and trade. *American Economic Journal: Applied Economics*, 14(1).



- Alfarano, S. and Milaković, M. (2008). Does classical competition explain the statistical features of firm growth? *Economics Letters*, 101(3):272 – 274.
- Amaral, L. A. N., Buldyrev, S. V., Havlin, S., Maass, P., Salinger, M. A., Stanley, H. E., and Stanley, M. H. (1997). Scaling behavior in economics: The problem of quantifying company growth. *Physica A: Statistical Mechanics and its Applications*, 244(1):1 – 24.
- Anderson, T. W. (1962). On the distribution of the two-sample Cramer-von Misés criterion. *Annals of Mathematical Statistics*, pages 1148–1159.
- Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23(2):193–212.
- Arrow, K. (1991). The dynamics of technological change. In *Technology and Productivity: The Challenge for Economic Policy*, pages 473–476. OECD Paris, France.
- Artzrouni, M. (1985). Generalized stable population theory. *Journal of Mathematical Biology*, 21(3):363–381.
- Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *Quarterly Journal of Economics*, 135(2):645–709.
- Axtell, R. L. (2001). Zipf distribution of US firm sizes. *Science*, 293(5536):1818–1820.
- Bargigli, L., Riccetti, L., Russo, A., and Gallegati, M. (2020). Network calibration and metamodeling of a financial accelerator agent based model. *Journal of Economic Interaction and Coordination*, 15:413–440.
- Böhm, F. (1960). Demokratie und unternehmerische Macht. In Institut für ausländisches und internationales Wirtschaftsrecht, editor, *Kartelle und Monopole im modernen Recht*, pages 1–24. C. F. Müller.
- Bottazzi, G. (2014). Subbotools user’s manual. *Laboratory of Economics and Management Working Paper Series, Sant’Anna School of Advanced Studies, Pisa (Italy)*, March 2014.
- Bottazzi, G., Cefis, E., and Dosi, G. (2002). Corporate growth and industrial structures: some evidence from the Italian manufacturing industry. *Industrial and Corporate Change*, 11(4):705–723.
- Bottazzi, G., Coad, A., Jacoby, N., and Secchi, A. (2011). Corporate growth and industrial dynamics: Evidence from French manufacturing. *Applied Economics*, 43(1):103–116.
- Bottazzi, G., Dosi, G., Lippi, M., Pammolli, F., and Riccaboni, M. (2001). Innovation and corporate growth in the evolution of the drug industry. *International Journal of Industrial Organization*, 19(7):1161–1187.
- Bottazzi, G. and Secchi, A. (2005). Growth and diversification patterns of the worldwide pharmaceutical industry. *Review of Industrial Organization*, 26(2):195–216.
- Bottazzi, G. and Secchi, A. (2006). Explaining the distribution of firm growth rates. *RAND Journal of Economics*, 37(2):235–256.

- Bray, M. M. and Savin, N. E. (1986). Rational expectations equilibria, learning, and model specification. *Econometrica*, 54(5):1129–1160.
- Brock, W. (1999). Scaling in economics: a reader’s guide. *Industrial and Corporate Change*, 8(3):409–446.
- Cantner, U. (2017). Foundations of economic change: An extended Schumpeterian approach. In Pyka, A. and Cantner, U., editors, *Foundations of Economic Change*, pages 9–49. Springer.
- Cantner, U. and Krüger, J. J. (2008). Micro-heterogeneity and aggregate productivity development in the German manufacturing sector. *Journal of Evolutionary Economics*, 18(2):119–133.
- Cantner, U., Krüger, J. J., and Söllner, R. (2012). Product quality, product price, and share dynamics in the German compact car market. *Industrial and Corporate Change*, 21(5):1085–1115.
- Cantner, U., Savin, I., and Vannuccini, S. (2019). Replicator dynamics in value chains: explaining some puzzles of market selection. *Industrial and Corporate Change*, 28(3):589–611.
- Chamberlin, E. H. (1949). *Theory of monopolistic competition: A re-orientation of the theory of value*. Oxford University Press, London.
- Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4):661–703.
- Coad, A. (2010). Investigating the exponential age distribution of firms. *Economics: The Open-Access, Open-Assessment E-Journal*, 4.
- Coad, A. and Tamvada, J. P. (2008). The growth and decline of small firms in developing countries. *Papers on Economics and Evolution*, 0808.
- Coleman, J. S. (1990). *Metatheory: Explanation in social science*, pages 1–23. The Belknap Press of Harvard University Press.
- Covarrubias, M., Gutiérrez, G., and Philippon, T. (2020). From good to bad concentration? US industries over the past 30 years. *NBER Macroeconomics Annual*, 34(1):1–46.
- Cunningham, C., Ederer, F., and Ma, S. (2021). Killer acquisitions. *Journal of Political Economy*, 129(3):649–702.
- Daepf, M. I., Hamilton, M. J., West, G. B., and Bettencourt, L. M. (2015). The mortality of companies. *Journal of The Royal Society Interface*, 12(106):1–8.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3):319–340.
- De Haan, L. and Resnick, S. (1998). On asymptotic normality of the Hill estimator. *Stochastic Models*, 14(4):849–866.
- Di Giovanni, J. and Levchenko, A. A. (2013). Firm entry, trade, and welfare in Zipf’s world. *Journal of International Economics*, 89(2):283–296.

- Di Giovanni, J., Levchenko, A. A., and Rancière, R. (2011). Power laws in firm size and openness to trade: Measurement and implications. *Journal of International Economics*, 85(1):42 – 52.
- Dold, M. and Krieger, T. (2019). *Ordoliberalism and European Economic Policy: Between Realpolitik and Economic Utopia*. Routledge.
- Dosi, G., Fagiolo, G., Napoletano, M., and Roventini, A. (2013). Income distribution, credit and fiscal policies in an agent-based Keynesian model. *Journal of Economic Dynamics and Control*, 37(8):1598 – 1625.
- Dosi, G., Fagiolo, G., Napoletano, M., Roventini, A., and Treibich, T. (2015). Fiscal and monetary policies in complex evolving economies. *Journal of Economic Dynamics and Control*, 52:166 – 189.
- Dosi, G., Fagiolo, G., and Roventini, A. (2010). Schumpeter meeting Keynes: A policy-friendly model of endogenous growth and business cycles. *Journal of Economic Dynamics and Control*, 34(9):1748–1767.
- Dosi, G., Marsili, O., Orsenigo, L., and Salvatore, R. (1995). Learning, market selection and the evolution of industrial structures. *Small Business Economics*, 7(6):411–436.
- Dosi, G., Napoletano, M., Roventini, A., and Treibich, T. (2017a). Micro and macro policies in the Keynes + Schumpeter evolutionary models. *Journal of Evolutionary Economics*, 27(1):63–90.
- Dosi, G., Pereira, M. C., and Virgillito, M. E. (2017b). The footprint of evolutionary processes of learning and selection upon the statistical properties of industrial dynamics. *Industrial and Corporate Change*, 26(2):187–210.
- Dosi, G., Pereira, M. C., and Virgillito, M. E. (2018). On the robustness of the fat-tailed distribution of firm growth rates: a global sensitivity analysis. *Journal of Economic Interaction and Coordination*, 13(1):173–193.
- Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, 4(5):41–60.
- Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5(1):17–60.
- Erlingsson, E. J., Alfarano, S., Raberto, M., and Stefánsson, H. (2013). On the distributional properties of size, profit and growth of Icelandic firms. *Journal of Economic Interaction and Coordination*, 8(1):57–74.
- Evans, G. W. and Honkapohja, S. (2013). Learning as a rational foundation for macroeconomics and finance. In Frydman, R. and Phelps, E. S., editors, *Rethinking expectations: The way forward for macroeconomics*. Princeton University Press.
- Fagiolo, G., Guerini, M., Lamperti, F., Moneta, A., and Roventini, A. (2019). Validation of agent-based models in economics and finance. In Beisbart, C. and Saam, N. J., editors, *Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives*, pages 763–787. Springer International Publishing, Cham.

- Fisher, R. (1930). *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- Fujiwara, Y., Di Guilmi, C., Aoyama, H., Gallegati, M., and Souma, W. (2004). Do Pareto–Zipf and Gibrat laws hold true? An analysis with European firms. *Physica A: Statistical Mechanics and its Applications*, 335(1-2):197–216.
- Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics*, 1(1):255–294.
- Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772.
- Gabaix, X. and Landier, A. (2008). Why has CEO pay increased so much? *Quarterly Journal of Economics*, 123(1):49–100.
- Geroski, P. A., Mata, J., and Portugal, P. (2010). Founding conditions and the survival of new firms. *Strategic Management Journal*, 31(5):510–529.
- Glauber, J., Wollersheim, J., Sandner, P., and Welp, I. M. (2015). The patenting activity of German universities. *Journal of Business Economics*, 85(7):719–757.
- Goldstein, M. L., Morris, S. A., and Yen, G. G. (2004). Problems with fitting to the power-law distribution. *European Physical Journal B*, 41(2):255–258.
- Gompers, P. A., Gornall, W., Kaplan, S. N., and Strebulaev, I. A. (2020). How do venture capitalists make decisions? *Journal of Financial Economics*, 135(1):169–190.
- Gourieroux, C. and Monfort, A. (1996). *Simulation-based econometric methods*. Oxford University Press.
- Greve, H. R. (2009). Bigger and safer: the diffusion of competitive advantage. *Strategic Management Journal*, 30(1):1–23.
- Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., and Railsback, S. F. (2010). The ODD protocol: a review and first update. *Ecological Modelling*, 221(23):2760–2768.
- Grüne-Yanoff, T. (2009). Learning from minimal economic models. *Erkenntnis*, 70(1):81–99.
- Grüne-Yanoff, T. and Verreault-Julien, P. (2021). How-possibly explanations in economics: anything goes? *Journal of Economic Methodology*, 28(1):114–123.
- Hassan, S., Pavón, J., Antunes, L., and Gilbert, N. (2010). Injecting data into agent-based simulation. In Takadama, K., Cioffi-Revilla, C., and Deffuant, G., editors, *Simulating Interacting Agents and Social Phenomena*, pages 177–191, Tokyo. Springer Japan.
- Hedström, P. and Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology*, 36(1):49–67.
- Hegselmann, R. and Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3).

- Heinrich, T. and Dai, S. (2016). Diversity of firm sizes, complexity, and industry structure in the Chinese economy. *Structural Change and Economic Dynamics*, 37:90–106.
- Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26):10570–10575.
- Johnson, S. L., Faraj, S., and Kudaravalli, S. (2014). Emergence of power laws in on-line communities: The role of social mechanisms and preferential attachment. *MIS Quarterly*, 38(3):795–808.
- Kaldor, N. (1961). Capital accumulation and economic growth. In *The Theory of Capital*, pages 177–222. Springer.
- Kang, S. H., Jiang, Z., Cheong, C., and Yoon, S.-M. (2011). Changes of firm size distribution: The case of Korea. *Physica A: Statistical Mechanics and its Applications*, 390(2):319–327.
- Kinsella, S. (2009). The age distribution of firms in Ireland, 1961–2009. Technical report, Department of Economics, University of Limerick, Ireland.
- Klein, D. and Marx, J. (2018). Wenn du gehst, geh ich auch! Die Rolle von Informationsskaskaden bei der Entstehung von Massenbewegungen. *PVS Politische Vierteljahresschrift*, 58(4):560–592.
- Klein, D., Marx, J., and Fischbach, K. (2018). Agent-Based Modeling in Social Science, History, and Philosophy. An Introduction. *Historical Social Research/Historische Sozialforschung*, 43(1 (163):7–27.
- Klein, D., Marx, J., and Scheller, S. (2019). Rational Choice and Asymmetric Learning in Iterated Social Interactions – Some Lessons from Agent-Based Modeling. In Marker, K., Schmitt, A., and Sirsch, J., editors, *Demokratie und Entscheidung Beiträge zur Analytischen Politischen Theorie*, pages 277–294. Springer.
- Kotz, S., Kozubowski, T., and Podgorski, K. (2012). *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance*. Springer Science & Business Media.
- Lehmann, E. L. and Casella, G. (2006). *Theory of point estimation*. Springer Science & Business Media.
- Luttmer, E. G. (2007). Selection, growth, and the size distribution of firms. *Quarterly Journal of Economics*, 122(3):1103–1144.
- Macy, M. W. and Flache, A. (2002). Learning dynamics in social dilemmas. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7229–7236.
- Mäki, U. (2009). MISSING the World. Models as Isolations and Credible Surrogate Systems. *Erkenntnis*, 70:29–43.
- Mayerhoffer, D. M. (2018). Raising children to be (in-)tolerant: Influence of church, education, and society on adolescents’ stance towards queer people in germany. *Historical Social Research*, 43(1):144–167.

- Mazzucato, M. (1998). A computational model of economies of scale and market share instability. *Structural Change and Economic Dynamics*, 9(1):55–83.
- Milani, F. (2007). Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics*, 54(7):2065–2082.
- Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., and Schwarz, N. (2013). Describing human decisions in agent-based models—ODD+ D, an extension of the ODD protocol. *Environmental Modelling & Software*, 48:37–48.
- Mundt, P., Alfarano, S., and Milaković, M. (2015). Gibrat’s law redux: think profitability instead of growth. *Industrial and Corporate Change*, 25(4):549–571.
- Newman, M. J. E. (2005). Power laws, Pareto distributions and Zipf’s law. *Contemporary Physics*, 4:323–351.
- Okuyama, K., Takayasu, M., and Takayasu, H. (1999). Zipf’s law in income distribution of companies. *Physica A: Statistical Mechanics and its Applications*, 269(1):125–131.
- Pascoal, R., Augusto, M., and Monteiro, A. M. (2016). Size distribution of Portuguese firms between 2006 and 2012. *Physica A: Statistical Mechanics and its Applications*, 458:342–355.
- Pellegrino, B. (2019). Product differentiation, oligopoly, and resource allocation. *WRDS Research Paper*.
- Robinson, J. (1969). *The economics of imperfect competition*. Springer.
- Ruppert, D. (2014). *Statistics and Finance: An introduction*. Springer.
- Schlag, K. H. (1998). Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78(1):130 – 156.
- Schlichte, F., Junge, S., and Mammen, J. (2019). Being at the right place at the right time: does the timing within technology waves determine new venture success? *Journal of Business Economics*, 89(8):995–1021.
- Silvestre, J. (1993). The market-power foundations of macroeconomic policy. *Journal of Economic Literature*, 31(1):105–141.
- Skilton, P. F. and Bernardes, E. (2015). Competition network structure and product market entry. *Strategic Management Journal*, 36(11):1688–1696.
- Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, 19(2):279–281.
- Sokol, D. D. (2014). Tensions between antitrust and industrial policy. *George Mason Law Review*, 22:1247.
- Stumpf, M. P. and Porter, M. A. (2012). Critical truths about power laws. *Science*, 335(6069):665–666.

- Subbotin, M. T. (1923). On the law of frequency of error. *Matematicheskii Sbornik*, 31(2):296–301.
- Sugden, R. (2009). Credible worlds, capacities and mechanisms. *Erkenntnis*, 70(1):3–27.
- Tesfatsion, L. (2017). Modeling economic systems as locally-constructive sequential games. *Journal of Economic Methodology*, 24(4):384–409.
- Wilensky, U. (1999). *NetLogo*. Northwestern University, Evanston.
- Will, O. and Hegselmann, R. (2014). Modeling Hume’s theory of norm emergence. In Conte, R., Andrighetto, G., and Campennl, M., editors, *Minding norms: Mechanisms and dynamics of social order in agent societies*, pages 68–80. Oxford Scholarship Online.
- Zhang, J., Chen, Q., and Wang, Y. (2009). Zipf distribution in top Chinese firms and an economic explanation. *Physica A: Statistical Mechanics and its Applications*, 388(10):2020–2024.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley Press, Boston.
- Zollman, K. J. S. (2009). The epistemic benefit of transient diversity. *Erkenntnis*, 72(1).

# A Network-Based Explanation of Inequality Perceptions

---

Joint work with A. Gebhard and D. M. Mayerhoffer.

Published in Social Networks (all rights reserved):

Schulz, J., Mayerhoffer, D. M., and Gebhard, A. (2022). A network-based explanation of inequality perceptions. *Social Networks*, 70:306–324. <https://doi.org/10.1016/j.socnet.2022.02.007>

We would like to express our gratitude to two anonymous reviewers and Martin Everett for his excellent editorial work, Arndt Leininger and his students, Miriam Bömer, Bettina Gregg, Johannes Marx, Moritz Schulz, Eleonora Priori and Jan Weber as well as the participants of the Networks 2021 conference, the 9th PhD conference for the renewal of constitutional economics, the 9th ECINEQ Meeting, the 33rd annual EAEPE conference, the annual PhD conference by the Hans Böckler Foundation 2021, the 2nd Scientific Workshop by the Network for Pluralist Economics, the 8th Workshop on Complexity, Innovation and Knowledge (WICK), the 14th Annual Conference of the International Network for Analytical Sociology, the 11th Young Economists Conference by the Vienna Chamber of Labour, the Research Seminar at the University of Utah and the 28th DVPW congress, especially our designated discussants Macartan Humphreys, Claudius Gräbner-Radkowsch and Pietro Terna which were of great help at crucial junctions of this investigation. Finally, we also thank Carsten Källner for his able research assistance. All remaining errors are, of course, ours.

---

## Abstract

Across income groups and countries, individual citizens perceive economic inequality spectacularly wrong. These misperceptions have far-reaching consequences, as it might be *perceived* inequality, not *actual* inequality informing redistributive preferences. The prevalence of this phenomenon is independent of social class and welfare regime, which suggests the existence of a common mechanism behind public perceptions. The literature has identified several stylised facts on how individual perceptions respond to actual inequality and how these biases vary systematically along the income distribution. We propose a network-based explanation of perceived inequality building on recent advances in random geometric graph theory. The generating mechanism can replicate all of aforementioned stylised facts simultaneously. It also produces social networks that exhibit salient features of real-world networks; namely, they cannot be statistically distinguished from small-world networks, testifying to the robustness of our approach. Our results, therefore, suggest that



homophilic segregation is a promising candidate to explain inequality perceptions with strong implications for theories of consumption and voting behaviour.

---

---

## Contents of this Paper

5.1	Introduction . . . . .	121
5.2	Related Literature . . . . .	123
5.2.1	Stylised Facts on Inequality Perceptions and Middle Class Bias . . .	124
5.2.2	Empirical Social Networks . . . . .	125
5.2.3	Belief Formation in Networks . . . . .	127
5.3	Model . . . . .	128
5.3.1	Agent Initialisation and Income Allocation . . . . .	129
5.3.2	Network Formation . . . . .	129
5.3.3	Gini Perception and Network Evaluation . . . . .	133
5.3.4	Validation . . . . .	134
5.4	Results . . . . .	135
5.4.1	Small-Worldiness . . . . .	136
5.4.2	Perceived Social Hierarchy and Middle Class Bias . . . . .	138
5.4.3	Perceived Individual Inequality . . . . .	140
5.4.4	Perceived Global Inequality . . . . .	142
5.4.5	Perception Dynamics . . . . .	143
5.4.6	Segregation Patterns . . . . .	144
5.4.7	Summary of Generating Mechanisms . . . . .	146
5.5	Discussion . . . . .	147
5.A	Linkage Probabilities for Homophilic Networks . . . . .	149
5.B	Perceived Quantiles in Perception Networks . . . . .	151
5.C	Linkage Based on Logarithmised Income Differences . . . . .	154
	References . . . . .	157

---

## 5.1 Introduction

Conventional modern macroeconomics has long recognised the crucial relevance of expectations and belief-formation for aggregate dynamics (Galí 2015). In particular, beliefs about economic inequality and perceptions of social hierarchy can inform individuals in such diverse fields as consumption decisions (Duesenberry 1949; Veblen 1899; Frank et al. 2014), redistributive preferences and voting behaviour (Gimpelson and Treisman 2018; Kim et al. 2018; Choi 2019) or subjective well-being and ethical convictions (Kuhn 2019; Clark and Senik 2010). Even in the most sophisticated behavioural models, belief-formation is, however, typically either assumed to be atomistic (Gabaix 2020) or does

not systematically account for the impact of individual embeddedness within heterogeneous social contexts on those beliefs, even if social interaction is explicitly modelled (Flieth and Foster 2002; Lux 2009). We propose a parsimonious network-based model for the interaction of macro-level inequality, micro-level beliefs and the mediating effects of heterogeneous social contexts. In contrast to the assumption of deductive reasoning in orthodox models, we build on the empirically well-established notion that economic agents reason *inductively* and generalise from finite samples. Recent theoretical and empirical work has demonstrated the potency of this approach in explaining phenomena in such diverse fields as human probability assessment (Sanborn and Chater 2016; Chater et al. 2020) or regional inequality (Collier and Tuckett 2020). The model is both consistent with several stylised facts about inequality perceptions and the micro-level evidence on the composition of social networks.

The relevance of individual beliefs is perhaps best exemplified by spelling out its political economy implications. Across income groups and countries, the public perception of economic inequality and many other macroeconomic variables is empirically wrong, often spectacularly so. Errors in those beliefs might be due to conceptually different problems: *uninformed* beliefs or *misinformed* beliefs (Kuklinski et al. 2000). Uninformed voters are ignorant about the actual state of affairs, while misinformed voters' beliefs are consistently deviating from it in one direction. The distinction is a crucial one. Uninformed voters' beliefs would cluster around the actual state of affairs and, with no systematic deviations, be correct in expectations. For uninformed voters, we only need one informed voter to tip elections under majority rule into the correct direction; a majority of ignorant individuals might nevertheless vote for the correct policy, which is now known as the 'miracle of aggregation' (Page and Shapiro 1993). However, this miraculous aggregation breaks down when we consider misinformed rather than uninformed voters with beliefs that are no longer randomly distributed but consistently tend in a (false) direction (Caplan 2011). The type of error in perceptions is thus intimately linked to the efficacy of democratic systems. For inequality perceptions, beliefs appear to be indeed the result of misinformation in this technical sense and they are consistently biased across income groups and welfare regimes.

In contrast to much of the behavioural literature, we refrain from ad-hoc assumptions about possible biases, e.g., assuming that individuals tend to perceive themselves in the middle of social hierarchies (cf., e.g. Knell and Stix 2020). Instead, we assume unbiased information processing capabilities for all the economic agents. Information is, however, asymmetric and agents form estimates about aggregate variables according to their local information. We show that a parsimonious process can generate sufficiently skewed information sets to replicate the aforementioned stylised facts and generate perceived inequality levels that are quantitatively in line with recent empirical evidence for a large sample of 32 OECD countries (Choi 2019). In essence, we assume that agents (correctly)

observe inequality within their local social network and (correctly) form estimates about the total population from them but still generate biased perceptions due to their network contacts not being representative for the overall population. Employing a new variant of a random geometric graph network, the assumption of income homophily alone can generate substantial misperception in line with the empirical evidence. The derived network topology also corresponds to empirically observed social networks across the world and features a small-world structure. Given the ubiquity of these topological features, our homophilic process appears to be a plausible candidate to explain the equally ubiquitous inequality misperceptions.

Our contribution is thus threefold: Firstly, we compile a list of four stylised empirical facts about income inequality and its biased perception from the nascent literature on the matter. Secondly, we develop a model that simultaneously replicates these stylised facts building on homophilic linkage and unbiased individual estimates based on local signals. Our model is quantitatively consistent with empirical estimates regarding both the input income distribution and the output perceptions, in contrast to the somewhat stylised models in the extant literature that also fail to replicate the dynamic behaviour of perceptions in response to changes in actual inequality. Thirdly, the network-formation algorithm presents a novel way of generating random geometric graph types of networks which is more intuitive for many application scenarios and allows specifying a minimum degree.

The remainder of this paper is organised as follows: Section 5.2 extracts four stylised empirical facts about inequality perception; furthermore, it reviews the evidence on empirical network topologies and individual belief formation within networks. Section 5.3 introduces the basic model of homophilic graph formation, reviews the main mechanisms generating heterogeneity in information sets, and validates the model. Section 5.4 presents our analytic and simulative results, shows that they are consistent with the outlined stylised facts regarding network topologies as well as inequality perceptions and derives some important implications regarding heterogeneous segregation patterns across the income distribution. Section 5.5 concludes and discusses several promising avenues for further research, especially regarding consumption dynamics and voting behaviour.

## 5.2 Related Literature

Our model joins three different strands of literature. Empirical findings on inequality perceptions that a single theory or model has not yet explained constitute its main explanandum. As explanans, we develop a network model featuring the current state of research into both the social network structure of empirical networks, mainly their homophily and small-world character, and individual perceptions in networks. The family

of random geometric graphs constitutes the third strand of literature as a promising methodological choice in Section 5.3.

### 5.2.1 Stylised Facts on Inequality Perceptions and Middle Class Bias

The empirical literature has identified four particular stylised facts for any theory of perceived inequality to be evaluated against: (i) Irrespective of their objective status, all individuals perceive themselves to be in the middle of the social hierarchy (Kelley and Evans 1995; Evans and Kelley 2004); (ii) as an immediate corollary of (i), poor individuals overestimate their social position, rich individuals tend to underestimate it (Knell and Stix 2020); (iii) poor individuals tend to perceive inequality to be higher and are closer to objective inequality on average (Osberg and Smeeding 2006; Newman et al. 2018) and (iv) the evolution of objective inequality is detached from the evolution of subjective inequality, that is, increases in objective inequality do not necessarily increase perceived inequality (Kenworthy and McCall 2008; Bartels 2018; Gimpelson and Treisman 2018; Hvidberg et al. 2020). The ubiquity of misperceptions across states and welfare regimes calls for a common mechanism independent of differences in actual inequality or institutional framework.

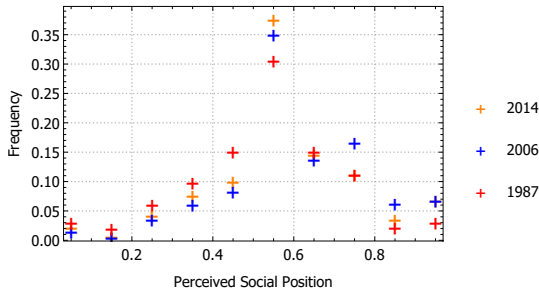


Figure 5.1: Empirical density of self-reported income decile by US respondents in the ISSP for the waves 1987, 2006, and 2014.

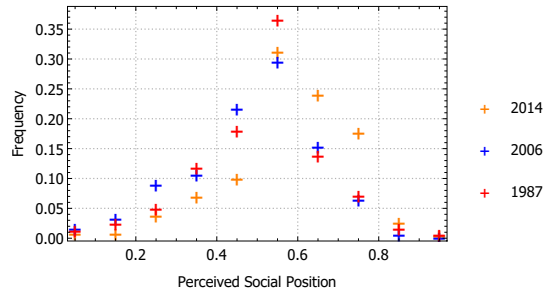


Figure 5.2: Empirical density of self-reported income decile by German respondents in the ISSP for the waves 1987, 2006, and 2014.

By way of illustration, we show the empirical frequency of self-perceptions to enable comparisons to the model output in Section 5.4 for stylised fact (i). The data shows the empirical frequency distributions of self-reported income deciles for Germany and the US from the 1987, 2007 and 2014 wave of the ISSP ISSP Research Group (2016).<sup>1</sup> Typically, Germany is considered to be the epitome of a *coordinated market economy*, while the US represents a *liberal market economy* (Hall and Soskice 2001). By contrast, the qualitatively similar densities for both countries show that the mechanism behind misperceptions

<sup>1</sup>We included exactly these years, since they represent the first, last and median waves of available data.

should be independent of the specific welfare regime. Especially the ‘middle class bias’ in the self-perceptions is rather striking. For a representative survey like the ISSP, each decile should, per construction, include exactly 10% of observations and the frequency density should therefore exhibit a uniform density at 0.1. Instead, the frequencies display a marked peak at the middle categories, i.e., most Germans and most Americans tend to think they are middle class, even though they are objectively not. This finding holds for all considered countries in the ISSP and all considered years, apart from the three periods and two countries we selected for illustrative purposes (Choi 2019).<sup>2</sup> Notice, however, that the qualitative middle class bias manifests itself in quantitatively rather different frequencies throughout time and between countries. We focus on the phenomenon that is common to all considered densities, i.e., the qualitative middle class bias and leave the direct quantitative calibration of our model for further research. We find this middle class bias to be the major driver of our results, implying the other stylised facts (ii) – (iv) directly but emerging endogenously in our network model. The relevant features of this network are discussed in the subsection below.

### 5.2.2 Empirical Social Networks

Empirical networks exhibit ubiquitous and salient features that can serve as stylised facts to guide the validation of proposed theoretical graph formation processes. Probably the most prominent one is the small-world property, indicating that paths between nodes in real-world social networks are unexpectedly short. At the same time, those networks also feature high degrees of clustering. Small-worldiness has obvious implications for any contagion process, be it rumours, diseases or information, where contagion across the whole network happens much faster than our intuition would suggest (Watts 1999; Moore and Newman 2000; Kleinberg 2001). The empirical research has identified small-world features across many different social groups, including friendship networks in schools (Weeden and Cornwell 2020), corporate board networks (Kogut and Walker 2001; Borgatti and Foster 2003; Davis et al. 2003; Conyon and Muldoon 2006; Galaskiewicz 2007) and scientific and artistic collaboration (Watts and Strogatz 1998; Newman 2001; Uzzi and Spiro 2005). Given this ubiquity, it appears safe to say that a graph-generating process for social networks needs to simultaneously produce low average path lengths and large degrees of clustering to be consistent with this stylised fact.

We situate our model in the random graph literature, where graph formation happens according to a stochastic process and is not the result of deliberate optimisation. Random graphs have been very successful in replicating structural stylised facts about network topologies, with the Watts-Strogatz model famously able to replicate those small-world properties (Watts and Strogatz 1998). Since the graph-generating process is, however,

---

<sup>2</sup>The other empirical frequencies are available upon request.

stochastic in nature, it does not feature explicit behavioural microfoundations. Therefore, we extend the purely stochastic notion with a behavioural ingredient, notably, that link-formation is *homophilic*.

First introduced by Lazarsfeld and Merton (1954), another salient feature of empirical social networks is *homophily*, the tendency of similar individuals to connect with each other. This tendency is not only an empirical curiosum but has relevant theoretical implications, e.g., for information transmission, where homophilic segregation can severely slow the speed of learning (Golub and Jackson 2012) or diminish the attention members of minority groups receive (Karimi et al. 2018). The similarity can come in many dimensions such as gender, ethnicity or socio-economic status (McPherson et al. 2001). We focus on the latter in the narrow sense of *economic* homophily, which is empirically well-established. One strand of literature focuses on friendship networks in schools and colleges and provides evidence significant homophily according to income or social class (Cohen 1979; Huckfeldt 1983; Mayer and Puller 2008; Boucher 2015; Malacarne 2017).

Even one of the earliest contributions in the field, however, shows that homophily in socio-economic status is not fixed in time and varies with cultural norms and the importance of class distinctions (Cohen 1979). In light of this result, it appears unsurprising that we find considerable variation in implied degrees of country-level homophily in our model, perhaps reflecting cultural norms not in the structure but the degree of the graph formation process. Even for a constant degree of (economic) homophily, increasing inequality also increases *segregation* in the population, as then the relative distances in incomes increase and agents becoming more selective in choosing links. For geographical segregation, this effect is empirically documented (Reardon and Bischoff 2011; Chen et al. 2012; Tóth et al. 2021), testifying indirectly to the relevance of homophily as a graph-generating feature. Homophily in income also exists in social media friendship networks (Lewis et al. 2012, for a large sample of Facebook friends), where spatial segregation should not confound findings and becomes apparent in the choice of romantic partners, typically under the label of ‘homogamy’ (Kalmijn 1991; Kalmijn and Flap 2001). Finally, a very recent contribution by Cepić and Tonković (2020) for a representative sample of Croatian adults finds evidence for homophilic tie formation according to social class and income, with however considerable variability in cross-class ties, hinting at possible confounding factors we aim to capture with a parsimonious stochastic process. Importantly, Cepić and Tonković (2020) show that there is also strong variation in cross-class ties, though, which we show might be crucial for unbiased individual inference. A major predictor is there shown to be cross-class sociability is political participation which would, in our model, amount to the prediction that politically active individuals are also better informed on average which is indeed perfectly in line with the evidence (Campbell 2013, for a recent survey).

### 5.2.3 Belief Formation in Networks

The literature on belief formation itself appears to be much more scarce than the literature on the effects of perceptions and misperceptions. While very different in detail, the two currently dominant theories of public opinion formation suggest that the beliefs an individual holds are in a broad sense averages over the idiosyncratic messages they receive (Zaller 1992; Lodge et al. 1995).<sup>3</sup> This literature has focused on the specific ‘averaging’ individuals use to process their information sets. Regarding perception formation about societal inequality, there exist to the best of our knowledge only two analytical models so far, namely Knell and Stix (2020) and Iacono and Ranaldi (2021). Both derive biased perceptions under very restrictive assumptions, though, and need to impose some kind of ‘biased averaging’. Knell and Stix (2020) assume that agents form subjective income densities over the whole support of possible income levels but let these densities be self-centered such that the mode of the distribution corresponds to their respective own incomes. While they partially succeed in replicating the four stylised facts on inequality perceptions at least qualitatively, their assumption essentially imposes the middle-class bias of stylised fact (i) *ex ante* and not as an emergent outcome from first principles. Iacono and Ranaldi (2021) also impose potentially biased perceptions and assume that agents only observe their own incomes as well as the minimum and maximum incomes. They continue to derive several important results on voting behaviour and show that information treatments on inequality might alter preferences for redistribution, therefore qualitatively replicating this stylised fact from the empirical literature. While the notion of local information sets appears appealing, using only the minimum, maximum and own income strikes us as unnecessarily artificial and implausible. Most importantly, being static, both models fail to make sense of the - arguably most relevant - stylised fact (iv), i.e., inequality perceptions being very persistent and not responding to changes in actual inequality. Our model is featuring this persistence by exploiting the fact that rising inequality also features rising segregation, as is also empirically established for geographical segregation (cf. the evidence discussed in subsection 5.2.2).

We develop on the notion that information is local but assume unbiased processing with skewed information sets as implied by the well-documented homophilic social network formation on which we expand in Section 5.3. The psychological literature on ‘social comparison theory’ (Festinger 1954) supports the notion that individual self-perceptions are much more responsive to local knowledge about small groups than to information about aggregates, e.g., knowing the population average (Buckingham and Alicke 2002; Zell and Alicke 2009; Alicke et al. 2010). Thus, belief formation about inequality appears to be indeed primarily based on local knowledge. This does not imply, however, that informational treatments in the form of reported averages do not change beliefs at all.

---

<sup>3</sup>Cf. Stevenson and Duch (2013) for a summary on those views.

Providing information about the actual degree of inequality seems to exhibit a significant effect on redistributive preferences for Argentina, Sweden and the US (Cruces et al. 2013; McCall et al. 2017; Karadja et al. 2017), with however small and insignificant effects for Germany (Engelhardt and Wagener 2018). Finally, two recent studies for the whole of Europe and Denmark separately demonstrate that individuals indeed tend to know the income levels of their immediate friends and family rather well, with non-negligible effects on inequality and fairness perceptions as well as perceived social positions (Clark and Senik 2010; Hvidberg et al. 2020).

Apart from this indirect evidence from informational treatments, there also exist several studies that measure the impact of local exposure to inequality on perceptions and redistributive preferences directly, therefore offering also direct evidence for our proposed mechanism.<sup>4</sup> Thal (2017) demonstrates using a large-scale survey (Soul of the Community (SOTC)) that affluent Americans' perception of social conditions is largely based on extrapolation from their own neighbourhood, as the affluent within homogeneous and isolated neighbourhoods perceive social disparities to be significantly less severe. Dawtry et al. (2015) find robust evidence for the US and New Zealand that individuals base their estimate of average societal income (and other quantiles) largely on their immediate subjective experience or 'social sample'. This leads to differences in fairness perceptions and redistributive preferences, whenever the composition of social circles varies by income, as homophily strongly indicates. Kraus et al. (2017) finally demonstrate with respect to racial economic inequality that the homogeneity of the immediate social network appears to mask racial inequities, therefore also testifying to the relevance of immediate lived experience for perception formation.

In a series of articles close in spirit to our approach, Chiang (2011, 2015a,b) exploits this notion and shows experimentally and computationally that individuals base their beliefs about inequality on local perceptions within referent networks and that income homophily has a potentially strong effect on those perceptions. While his approach is exploratory and does not account for the outlined stylised facts on inequality perceptions and empirical social networks, we provide a tractable model, readily calibrated with regards to those phenomena that is introduced below.

### 5.3 Model

This section provides a content-oriented presentation; a technical description following the ODD protocol is available upon request. A commented version of the model can be found at <https://github.com/mayerhoffer/Inequality-Perception>. The model consists of three distinct phases run in sequential order:

---

<sup>4</sup>We thank an anonymous reviewer for this hint and the helpful references.



1. Agent initialisation and income allocation
2. Network formation
3. Gini perception and network evaluation

Each phase runs only once and phases one and two build the structure which phase three then analyses. This sequence implies that during network generation, agents adapt to others' income level. However, there is no reaction to others' linking behaviour or perception and, thus, the model does not feature interaction in a narrow sense. Moreover, in the model, an agent's social contacts depend on their income. We choose this direction of causality for technical reasons and because it seems empirically likely (cf. Section 5.2). Nevertheless, our process scheduling would also be consistent with the opposite direction of causality or positive feedback effects between income and social contacts.

The model is designed that way because it focuses entirely on income perceptions given defined income distributions and network structures. Hence, both an agent's income and their social contacts remain constant for the evaluated time frame or, put differently, that the simulation outcome is a snapshot of a certain point in time.

### 5.3.1 Agent Initialisation and Income Allocation

There are 1,000 agents in the model; each agent draws their income from an exponential distribution with a mean of  $\lambda = 1$ . Such a distribution normalises the empirical observed (pre-tax or market) income distributions in various industrialised countries for the vast majority of individuals (Drăgulescu and Yakovenko 2001; Silva and Yakovenko 2004; Tao et al. 2019). Thus, one can understand the model population as constituting a representative sample of empirical populations of these countries. The upper tail of 1 to 5 % of the income distributions empirically follows a Pareto law (Silva and Yakovenko 2004). We deliberately choose to exclude this small minority from our model, since their population size would induce another degree of freedom in our model and we want to demonstrate that segregation is indeed endogenous and not driven by differences in actual income regime. We use an identical, pre-validated exponential distribution for all Monte Carlo runs and also all levels of homophily to ensure comparability between simulation runs. Agents store their true income decile for evaluation purposes, too.

### 5.3.2 Network Formation

Each agent draws five other agents to link to. Like for real-world networks, links are therefore created *by agents*, not imposed on them. The number of five link choices is also empirically validated, as humans tend to only know the income of close friends or family (Clark and Senik 2010; Hvidberg et al. 2020), with typically only five individuals at this

closest layer of emotional connection (Zhou et al. 2005; Hamilton et al. 2007; MacCarron et al. 2016).<sup>5</sup> The relative weight in the draws are a function of the homophily strength and the respective income levels. Thereby, agent  $j$ 's weight in agent  $i$ 's draw is denoted by  $w_{ij}$  and determined as follows:

$$w_{ij} = \frac{1}{\exp[\rho |I_j - I_i|]} \quad (5.1)$$

$I$  denotes the income of an agent, and  $\rho \in \mathbb{R}^+$  denotes the homophily strength in income selection, externally set, and identical for all agents.  $\rho = 0$  represents a random graph, and for an increasing positive value of  $\rho$ , an agent becomes ever more likely to pick link-neighbours with incomes being closer to their own. The exponential character of the link function ensures that those others with are large income difference become unlikely picks even at low homophily strengths.

The choice of an exponential weighting function might seem arbitrary but upon closer inspection, we find that translated into the probability of  $i$  choosing  $j$ , this weighting is equivalent to the discrete choice approach developed and popularised by Manski and McFadden (1981). The homophily parameter  $\rho \in (0, \infty)$  is then simply the intensity of choice parameter. To translate weights into probabilities, we normalise by all weights for all agents, i.e.,

$$p_{ij} = \frac{\exp[-\rho \cdot |I_j - I_i|]}{\sum_{k \in M \setminus i} \exp[-\rho \cdot |I_k - I_i|]}, \quad (5.2)$$

with  $M \setminus i$  as the set of all agents except  $i$  with size  $N - 1$ .<sup>6</sup> This formulation in (5.2) has a rather intuitive interpretation, with  $\rho = 0$  implying equiprobable picks with  $p_{ij} = 1/(N - 1)$ ,  $\forall j \in M \setminus i$ , and thus indeed a random graph, while  $\rho \rightarrow \infty$  implies that  $p$  approaches unity for  $j$  with minimum income distance and 0 for all other  $j$ . Manski and McFadden (1981) demonstrate that the discrete choice rule above emerges naturally from random utility theory, i.e., agents maximise utility and utility can be decomposed into an observable and unobservable component. In our case, the observable component the agents minimise would be the income differences, with the unobservable part being all the attributes from which our agent in question would benefit due to their social connection. This appears to be rather intuitive, since of course income differences might be a rather

---

<sup>5</sup>For a recent review on the large literature on ‘Dunbar’s number’, cf. the first section of MacCarron et al. (2016). However, to the best of our knowledge, there are no empirical studies that specify whom this closest layer consists of.

<sup>6</sup>Note that this is, strictly speaking, only the probability of the first choice of agent  $i$ , since we consider drawing without replacement, and does not account for the possibility that other agents already link to the agent in question, in contrast to our algorithm. Since the number of agents is rather large, the effect appears to cancel out in the aggregate, though, as also our simulation results in subsection 5.4.6 indicate.

salient characteristic and thus observable, while the utility from social connections might in some cases plausibly exceed the one derived from merely a good fit or small social distance.<sup>7</sup> In this sense, the weighting function in eq. (5.1) is plausibly microfounded in a utility-maximising framework and can now be considered the workhorse choice rule in behavioural macroeconomics (Franke and Westerhoff 2017). Franke and Westerhoff (2017) also survey evidence from a several lab experiments in different macroeconomic contexts that discrete choice is indeed consistent with the data, while Anufriev and Hommes (2012) and Anufriev et al. (2018) provide laboratory evidence for the discrete choice approach for financial markets. However, there might of course be other potential choice mechanisms that could provide avenues for further research on network generation that can be readily included within our proposed flexible RGG framework.

Figures 5.3 and 5.4 illustrate the linkage probabilities implied by the weighted draw based on the exponentially distributed income levels. As can be seen, the decay within the left tail is always more rapid than for the right tail, indicating differences in the ‘selectivity’ above or below a relative position. We understand ‘selectivity’ according to rank as the effect a decrease in income rank distance of one agent to another has on the linkage probability between them. Consequently, the local maxima of individual linkage probability densities exhibit a bi-modal shape with peaks at the highest and lowest rank but are also heavily skewed to the left, i.e., agents with the high incomes are most selective in their link picks. General selectivity increases with  $\rho$ . Notice also that largest income ranks are extremely selective in all scenarios, in some cases exceeding linkage probabilities of incomes close to the median by more than two orders of magnitude in linkage probabilities.

An anonymous reviewer pointed out that individuals may form links based on relative rather than absolute income differences: For lower incomes, a given absolute gap in units of currency may mean two entirely different lived experiences, while people with high incomes may hardly notice the same absolute gap. To represent this in the linkage function in eq. (5.1), one must simply replace  $I_i$  by  $\ln(I_i)$  and  $I_j$  by  $\ln(I_j)$ . 5.C analyses this transformation of scale in detail. The altered argument in the choice function is equivalent to assuming that agents aim to minimise the percentage difference in incomes and is, therefore, a natural extension to capture potentially scale-dependent tie-formation along the lines discussed above (Törnqvist et al. 1985).

Our findings from Section 5.4 regarding self-perceptions and aggregate inequality perceptions prove qualitatively robust; quantitatively, the major findings occur at even lower

---

<sup>7</sup>Notice, however, that the derivation of the above choice rule crucially depends on the axiom of Independence of Irrelevant Alternatives (IIA) (Luce 1977), i.e., the probability of choosing between  $j$  and  $k$  being independent of the probability of choosing  $l$ . IIA might be a good first-order approximation for homophilic choice but in friendship networks, knowing one agent  $j$  might indeed increase the likelihood of knowing another agent  $l$  that is friends with  $j$ . It might thus prove interesting to extend and generalise the above choice rule to examine the effects on the network topology in further research.

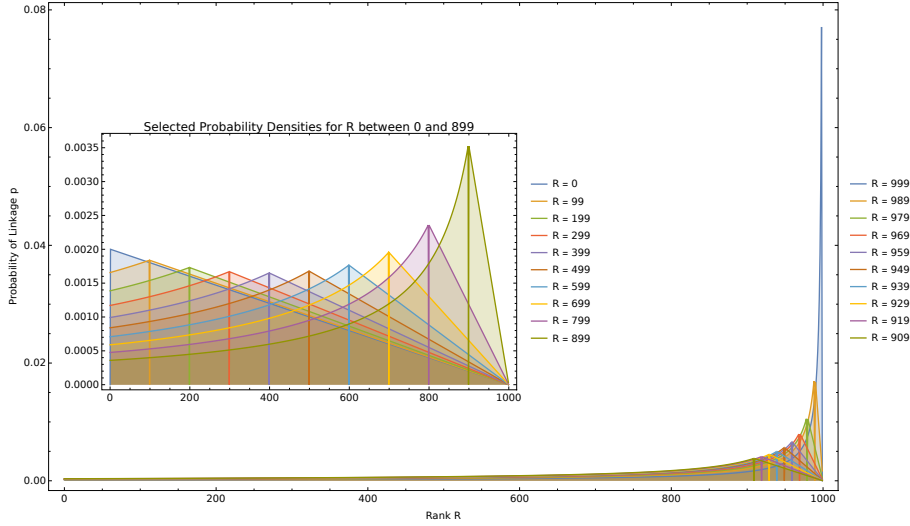


Figure 5.3: Theoretical PDF of Linkage Probabilities for Ranks  $R$  and  $\rho = 1$ .

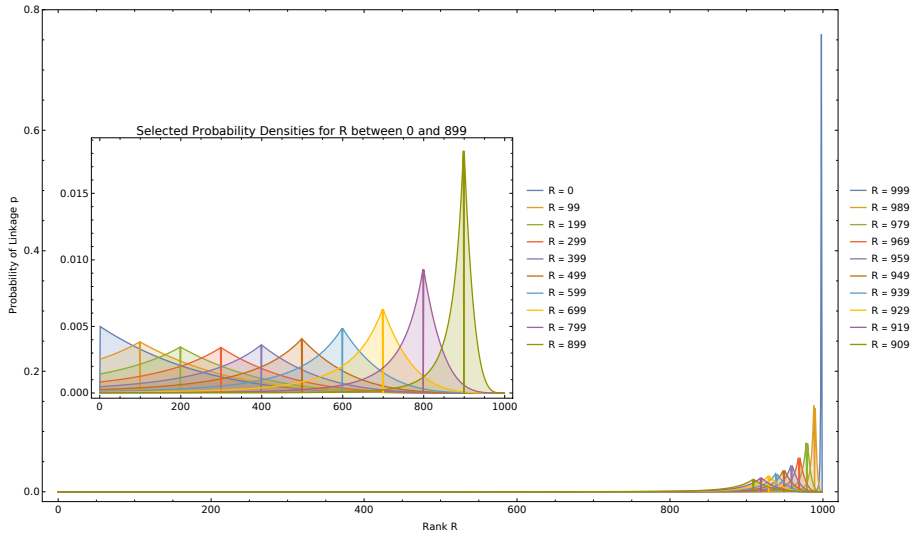


Figure 5.4: Theoretical PDF of Linkage Probabilities for Ranks  $R$  and  $\rho = 4$ .

*Note:* The Figures plot the Probability Density Functions (PDFs) of a node with a given income rank for linkage with another node for the whole support of income ranks. The above Figure assumes a homophily strength  $\rho = 1$ , whereas the below Figure assumes  $\rho = 4$ .

homophily levels. However, the logarithmised incomes fail to replicate the greater underestimation of inequality for richer individuals with higher income ranks. The segregation tendency is approximately symmetric for moderate to high homophily strengths, i.e., all agents are approximately equally likely to include agents below and above them in income rank. Since there is hardly any differential behaviour according to income rank, all agents tend to perceive roughly equal levels of inequality according to the (local) Gini, in contrast to the stylised fact (iii). Thus, absolute income differences pose the more

strongly validated presentation in the present model framework. Nevertheless, the model invites empirical research into whether income homophily is based on absolute or relative differences - or some combination of both.<sup>8</sup>

The resulting network for our baseline specification is a member of the family of Random Geometric Graphs (Dall and Christensen 2002), which Talaga and Nowak (2020) showed to reproduce core features of many social networks efficiently. Specifically, we combine the notions of homophily (Boguná et al. 2004) with pre-setting node degrees (Newman et al. 2001; Newman 2009).<sup>9</sup> However, concerning our application, we are able to simplify both approaches by pre-determination of only the global minimum degree, like in Preferential-Attachment networks, and consequently defining relative weights rather than absolute probabilities.

Links are undirected and have identical weights for evaluation purposes. Agents pick their neighbours in random sequential order. If an agent  $i$  picks agent  $j$  who had themselves picked  $i$  before that, the already existing link between the two agents remains untouched, but  $i$  does not pick another neighbour instead of  $j$ . Consequently, each agent has at least 5 link-neighbours (i.e. close social contacts with mutual knowledge of income) but may have more.

### 5.3.3 Gini Perception and Network Evaluation

Agents know about their own income and also their social contacts' incomes. However, they do not possess knowledge about any other agent or structural features of the whole income distribution. Thus, agents judge income inequality in the population as well as their own income position solely based on themselves and their link-neighbours. Besides the agents' perceptions, there is a global assessment of various network parameters in order to validate the model.

Subjective inequality perceptions mirror standard Gini calculation on the level of individual personal networks: Each agent finds the mean of all income differences between themselves and each link neighbour and between any two of their link-neighbours and divides this by twice the mean overall income of themselves and all link-neighbours. Then, the overall perceived Gini is simply the arithmetic mean of individual perceptions.

To estimate their income decile, an agent compares the number of link-neighbours having a higher income than the agent themselves to the link-neighbours having a lower income than the agent themselves.

---

<sup>8</sup>Since only the specification following scale-independent choice is consistent with inequality perceptions varying systematically with income while choice based on logarithmised incomes is not, the existence or non-existence of differential behaviour might help to discriminate between implied choice functions.

<sup>9</sup>Our procedure is related to Karimi et al. (2018) who also combine a Preferential-Attachment model with a homophilic ingredient. Crucially, however, their model builds on a binary notion of homophily with only two groups. Our algorithm, in contrast, does not impose any restriction on the target feature and is applicable to attributes potentially defined over the whole positive real half-line.

### 5.3.4 Validation

The simulation results of our model are in line with our theoretical expectations and we can explain their emergence in terms of the mechanisms sketched in Section 5.4. Moreover, we carried out sensitivity analyses that revealed no unintended consequences of changes in any relevant model feature like homophily level, number of links or actual income distribution. Thus, we consider the model design and implementation to be internally validated (Gilbert and Troitzsch 2005, p. 22) as a tool for explaining inequality perceptions in the model population.

Transferring these explanations from the model to the real world requires external validation of our model. However, there are different accounts of what constitutes an explanation in the first place. The current discussion of the concept of explanations in the philosophy of the social sciences highlights two types of explanations: How-actually and how-possibly explanations, also known as candidate explanations (Epstein 1999). While how-actually explanations aim for identifying the actual mechanism driving the dynamics in a specific case, how-possibly-explanations provide mechanisms that could possibly bring about the explanandum in question (Reutlinger et al. 2018); they enquire for mechanisms that potentially cause the observed phenomenon. In case of epistemically possible how-possibly explanations, these mechanisms are in line with the knowledge about the real world (Grüne-Yanoff and Verreault-Julien 2021).

Our model yields an epistemically possible how-possibly explanation of inequality perception because it “produces quantitative agreement with empirical macrostructures, as established through on-board statistical estimation routines” and also “quantitative agreement with empirical microstructures, as determined from cross-sectional and longitudinal analysis of the agent population” (Barde and Van Der Hoog 2017): Simulation outputs of a societal structure close to a small-world one with self-segregation of highest-income agents and severe underestimation of the income Gini across income levels mirror the corresponding empirical findings.

Following the suggestion by Fagiolo et al. (2019), we use empirical micro-data to calibrate the model. Namely, it relies on an exponential income distribution that characterises industrialised countries. Furthermore, the extent of agents’ closest layer of interaction (‘Dunbar’s number’) that means mutual knowledge of income, their linking behaviour, and individual perception formation follows rules that are theoretically established in rational choice theory but also empirically grounded in the referenced lab experiments and surveys. The exponential weighting function from the discrete choice framework is also analytically convenient and lets us represent the probability densities of ties in closed form. This allows us to e.g. demonstrate conclusively that the combination of discrete choice in graph formation and an exponential income distribution leads to the

endogenous emergence of echo chambers for top-income earners whose isolation increases in the intensity of choice  $\rho$ .<sup>10</sup>

This empirical input calibration and output validation jointly guarantee resemblance (Mäki 2009) between our model and the real world. We develop a specific parallel reality (Sugden 2009) that features generating mechanisms for empirical findings in our reality, and hence our results present a candidate explanation for the stylised empirical facts. There may be different, more adequate, parallel realities featuring either these or even better mechanisms, despite to the best of our knowledge there being no existing models that fulfil these characteristics. Overall, the following section presents an epistemically possible how-possibly explanation of inequality underestimation that “constitutes epistemic progress on the way towards HAEs [how-actually explanations, A/N]” (Grüne-Yanoff and Verreault-Julien 2021) of the phenomenon. The model simultaneously features technical verification and external validation based on input and output measures. Gräbner (2018) considers this combination desirable albeit rarely possible for model development. Since our model features a range of proposed micro-mechanisms (e.g., on endogenously evolving segregation, cf. subsection 5.4.5), we also hope to inform empirical research to further examine their external validity.

## 5.4 Results

The homophilic graph model will be evaluated against the five stylised facts outlined earlier. As we have shown in Section 5.3, we only require the homophily strength parameter  $\rho \in \mathbb{R}_0^+$ , the number of links each node chooses  $C$  and the income distribution as inputs for initialisation. Since link formation is stochastic, we run the graph formation routine 100 times and report model averages, if not otherwise indicated. Most of the results are obtained with initialisation by the same set of incomes generated from an exponential distribution with location parameter  $\lambda = 1$  and 1,000 observations for  $C = 5$  choices of link-neighbours each agent undertakes to make results comparable for variation in  $\rho$ .

---

<sup>10</sup>This combination of analytical convenience that leads to internal validity and empirical plausibility that affirms external validity is also one of the reasons why we deliberately choose not to use an Exponential Random Graph (ERGM) or Stochastic Actor Oriented Modelling (SAOM) framework (Snijders 2011, for a recent survey) but situate our model in the RGG framework: Firstly, the application of these types of models would require merging relational data with the socioeconomic status of the respective agents which is rarely achieved in practice, as De Paula (2017) notes. In our case, the problem of data availability is compounded by the fact that we require the graph data not only to report all social ties but also to identify the closest layer of emotional connection. Only there we can reasonably expect agents to exactly observe incomes as is required by our model mechanism. We are currently not aware of any dataset fulfilling these constraints but welcome any empirical attempt in this direction, as the external validity of our proposed model mechanism can ultimately only be established empirically. Secondly and more importantly, the estimated coefficient estimates and tie-level probability densities from ERGMs and SAOMs are purely phenomenological and need to be simulated by Monte Carlo techniques, while we are able to express them analytically and thus precisely determine the effect of our model parameters. We thank an anonymous reviewer for pointing us to ERGMs.

The overall Gini coefficient for these 1,000 randomly generated income levels is with  $G \approx 0.50701$  within 1.5 % deviation from the theoretical Gini of  $G = 0.5$ , indicating that the observed effects of  $\rho$  are not artefacts of initialisation. Results are also robust for different numbers of links chosen per node, as long as  $C \ll N$ . We also evaluated the null model for  $\rho = 0$ , where we did not find any significant deviations in the mean inequality perceptions and the actual overall inequality of  $G = 0.5$ , testifying to the robustness of our approach.<sup>11</sup>

### 5.4.1 Small-Worldiness

We use state-of-the-art methods to test for the existence of small-world features against an appropriate network null model, here an Erdős-Rényi (ER) graph with the corresponding number of nodes and mean degree first introduced by Erdős and Rényi (1960). ER graphs appear to be the correct null model for two reasons: Firstly, they are a particular case of our model with  $\rho = 0$ , i.e. without homophily. Hence, the procedure allows isolating the impact of homophily and examining whether the model indeed tends to yield ‘smaller worlds’ for homophilic formation in the precise sense outlined below. Secondly, we can establish an exact one-to-one correspondence between a graph generated by our model and the ER model, as ER graphs only require the number of nodes and a linkage probability for initialisation that is fully determined by the mean degree of the correspondent network. Other prominent generating models such as Watts-Strogatz graphs have additional degrees of freedom like the ‘rewiring probability’ without clear correspondence to our model.

We construct three summary metrics to test our model against, as introduced by Humphries and Gurney (2008). Firstly,  $\Lambda$  measures the deviation in average path lengths  $L$ , that is,

$$\Lambda_i := \frac{L_i}{L_i^{ER}}, \quad (5.3)$$

where  $L_i$  is the average path length of network  $i$  with  $L_i^{ER}$  as the average path length of a correspondent ER graph with equivalent number of nodes and mean degree. ‘Small-worldiness’ requires  $\Lambda \approx 1$ , as our network should not deviate too much from the random benchmark that indeed features short paths.  $E[L_i^{ER}] = (\log[N] - \gamma)/(\log[k]) + 1/2$  with  $\gamma$  as Euler’s constant,  $N$  as the number of nodes and  $k$  as the average degree can be analytically derived which we use in our calculation (Fronczak et al. 2004).

Secondly, we also require a high clustering coefficient which an ER graph cannot generate. The deviation in the clustering coefficients  $\Gamma$  is defined as

---

<sup>11</sup>The results for the null model as well as for different  $C$  are available upon request.



$$\Gamma_i := \frac{C_i}{C_i^{ER}}, \quad (5.4)$$

with  $C_i$  as the clustering coefficient of graph  $i$  and  $C_i^{ER}$  as the clustering coefficient of the corresponding ER graph. Here, again, analytical results are available which we utilise, mainly that  $E[C_i^{ER}] = k/N$  with again  $k$  as the average degree and  $N$  the number of nodes (Watts 1999). Since ER graphs typically do not exhibit clustering, we require here that  $\Gamma_i > 1$  for a small-world to be present.

Finally, we use a summary measure  $\Phi$  introduced by Humphries and Gurney (2008). We define  $\Phi$  as

$$\Phi_i := \frac{C_i}{C_i^{ER}} / \frac{L_i}{L_i^{ER}} = \frac{\Gamma_i}{\Lambda_i}. \quad (5.5)$$

Humphries and Gurney (2008) show that  $\Phi_i$  features desirable statistical properties when confronted with the conventional Watts-Strogatz model for graph formation and shows a unique maximum between the extreme cases of a random network and an ordered lattice. This is in line with our intuition that small-worldiness results from the interaction of order (in the form of high clustering near the lattice) and randomness (in the form of the random graph featuring low average path lengths), as shown by Watts and Strogatz (1998). We require  $\Phi > 1$  for small-worlds. Note that  $\Phi > 1$  is an immediate corollary of the two requirements  $\Gamma > 1$  and  $\Lambda \approx 1$ , but  $\Phi > 1$  does not imply the two individual requirements. We call the first sufficient condition ‘strong small-worldiness’ and  $\Phi > 1$  with a violation of either  $\Gamma > 1$  or  $\Lambda \approx 1$  ‘weak small-worldiness’, where we now only require normalised clustering to increase *faster* than average path lengths.

The ER benchmark is nested in our model for  $\rho = 0$ , as is also readily visible from the fact that both  $\Lambda \approx 1$  and  $\Gamma \approx 1$  for  $\rho = 0$ . We indeed find that homophily induces path lengths to grow significantly above this ER benchmark. Normalised clustering coefficients, however, increase much more rapidly with homophily than average path lengths, demonstrating that our model can achieve relatively high clustering without simultaneously increasing path lengths in the same way. The proposed process thus violates the strong condition but fulfills the weak condition for small-worlds and is therefore broadly in accordance with the topological patterns found in real-world social networks. We note further the symmetry to the canonical Watts-Strogatz approach (Watts and Strogatz 1998). While we build on a random network with short average path lengths and interpolate to the desired high clustering through homophily, Watts and Strogatz start from an ordered state with high clustering and approach the random graph benchmark by rewiring to generate shorter average path lengths. Arguably, however, our approach

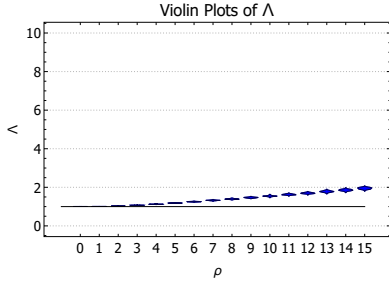


Figure 5.5: Violin Plots of Normalised Average Path Lengths  $\Lambda$  as a function of Homophily Strength  $\rho$ .

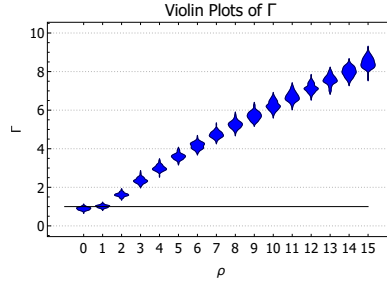


Figure 5.6: Violin Plots of Normalised Clustering Coefficients  $\Gamma$  as a function of Homophily Strength  $\rho$ .

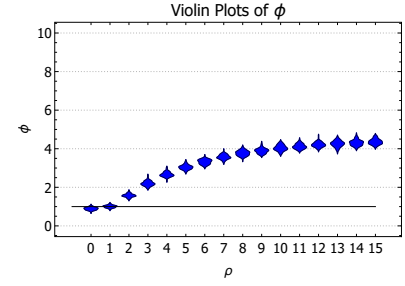


Figure 5.7: Violin Plots of Small-World Summary Statistic  $\Phi$  as a function of Homophily Strength  $\rho$ .

*Note:* The Figures report violin plots for the relevant statistics for ‘small-worldiness’. The average path length is significantly higher than the ER benchmark for all depicted  $\rho$ , indicating violation of the ‘strong small-worldiness condition’. Normalised clustering coefficients are for  $\rho > 1$  significantly higher than the ER benchmark and increase at a much faster rate than average path lengths, indicating that indeed the ‘weak small-worldiness’ condition is fulfilled.

starts from a plausible and empirically well-established behavioural principle in contrast to the purely stochastic process in the Watts-Strogatz world without such behavioural foundations. Besides providing empirical validation, this finding might also point to relatively rapid contagion throughout the homophilic network, be it in the form of rumours or ‘expenditure cascades’.

### 5.4.2 Perceived Social Hierarchy and Middle Class Bias

For unbiased hierarchy perceptions, the reported frequency of perceived social position would coincide with the actual positions. Unbiased perceptions thus entail reported perceived positions of equal frequency, as they coincide with the actual population shares. As we show both analytically in 5.B and by simulation, perceived social positions for homophilic graph formation are far from the equiprobable benchmark. We find a tendency of the vast majority of individuals to place themselves in the middle of the perceived hierarchy, in line with the empirical evidence. We prove that the tendency exists for all  $\rho \in (0, \infty)$ . Its strength is a function of  $\rho$ , though, as we show exemplarily in Figures 5.8 to 5.11. The figures plot the empirical densities of income ranks which the respective individuals perceive to hold. For  $\rho = 1$ , the tendency is relatively weak, while for  $\rho = 4$ ,  $\rho = 8$  and  $\rho = 14$ , the densities display a distribution that notably peaks at the centre. In fact, the displayed densities indeed seem to feature all the salient features of the densities of empirical perceived social positions, as shown in Choi (2019) and also in Figures 5.1 and 5.2 for the ISSP data we collect.

Notice that this a necessary *outcome* of homophilic graph formation under very mild and general conditions and based on a well-established utility maximisation framework, in contrast to models that take this tendency as an *assumption*. The latter strand of lit-

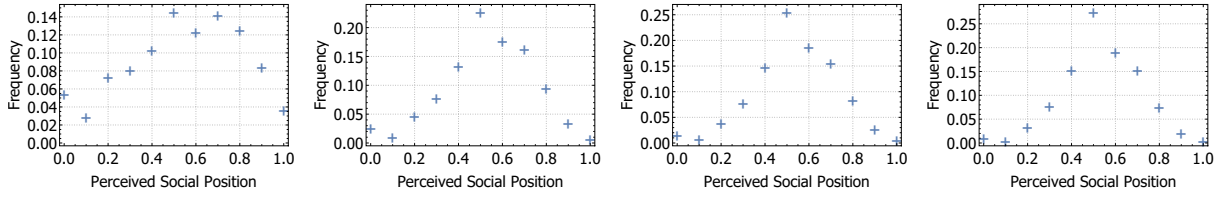


Figure 5.8: Density of Perceived Quantiles for  $\rho = 1$ . Figure 5.9: Density of Perceived Quantiles for  $\rho = 4$ . Figure 5.10: Density of Perceived Quantiles for  $\rho = 8$ . Figure 5.11: Density of Perceived Quantiles for  $\rho = 14$ .

*Note:* The Figures report the perceived social positions for  $\rho \in \{1; 4; 8; 14\}$  with 10 bins each. All Figures exhibit significant deviation from the benchmark with equal frequencies. The tendency for individuals to place themselves in the middle of the income hierarchy is, however, only apparent for the middle and right panels, indicating that a homophily strength  $\rho$  of 1 might be too low to replicate the empirically observed tendency. For  $\rho = 4, 8$  and  $14$ , the densities approximate the empirical densities rather well, though.

erature has typically taken a bounded rationality stance on the issue and argued that it is failures in information processing which explain the persistent errors in perceptions of social positioning. Our model replicates stylised fact (i) purely by virtue of the network formation process. In contrast to the literature on bounded rationality, we can hence show that stylised fact (i) is consistent with purely rational actors that form correct beliefs based on their available information, as long as homophilic graph formation constrains their information sets. Our model thus entails very different policy implications to improve self-perceptions. Since information processing is assumed to be correct in our model, information treatments, i.e., increasing the information received from nodes with heterophilic incomes, have mitigating effects on perceptions. Influencing information processing itself, as implied by the established models, is arguably a much harder task for policy.

An immediate corollary of the population perceiving themselves to earn the median income is the tendency for rather poor individuals to overestimate their position and the rich to underestimate it, as all perceive themselves to be in the middle. Hence, the fit for the median perception tracks the trend in the simulations reasonably well for the vast majority of observations (cf. Figure 5.12). As we discuss in more detail in 5.B, there is no tendency to the median for the left and right tail of the distributions which the simulation results reflect, too. Indeed, approaching the minimum or maximum improves the accuracy of individual estimates. The intuition for this is quite simple: The poorest and the richest individual will always correctly perceive their social position, independent of  $\rho \in \mathbb{R}_0^+$ . The rationale for this is that the actual minimum (maximum) of the whole will always be the minimum (maximum) of any potential non-empty subset of the population. Apart from such boundary effects, however, we indeed replicate stylised fact (ii) insofar as the poorer half of the population seems to overestimate their social position, while the richer half underestimates it. This finding is in line with the empirical evidence and suggests that

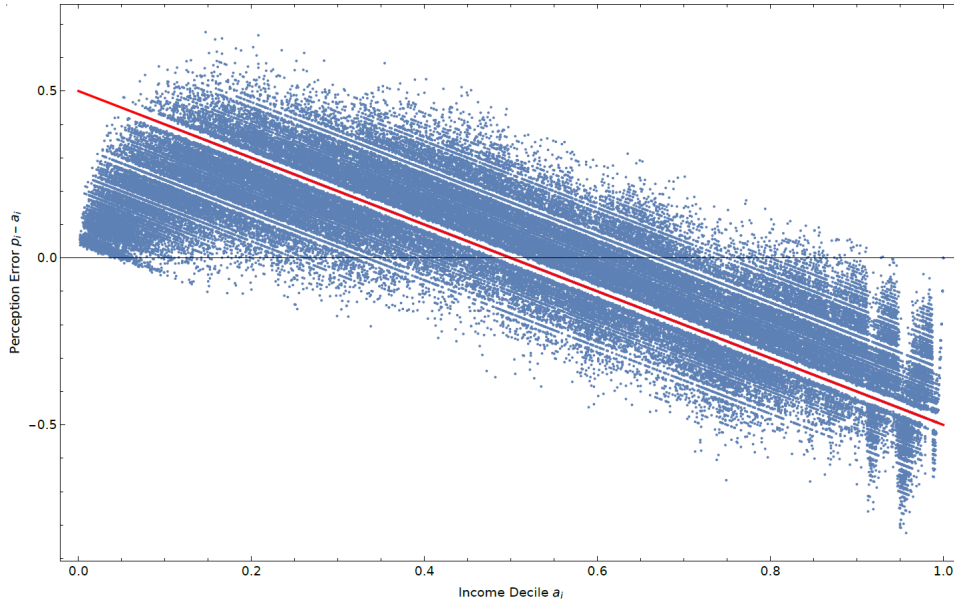


Figure 5.12: Errors  $\epsilon_i$  show the difference between perceived position  $q_i$  and actual position  $a_i$  for all individuals  $i$ , and  $\rho = 4$ . The superposed line corresponds to  $\epsilon_i = 0.5 - a_i$  or the belief for all individuals to be in a median position of the income distribution. Except for the boundary regions close to the minimum and maximum income, the theoretical fit approximates the trend in the data reasonably well. This indicates that the trend to the median is indeed present for the vast majority of the population.

total whole population tends to underestimate the degree of inequality, as we will show in the upcoming subsection.

### 5.4.3 Perceived Individual Inequality

We define perceived inequality as the Gini coefficient calculated over the perception set of a given individual  $i$ . In Figure 5.13, we plot those perceived Ginis against the income ranks of our individuals with a higher rank indicating a higher income. In line with stylised fact (iii), we find that inequality perceptions decrease approximately monotonically in income rank, while almost all individuals underestimate the actual degree of inequality significantly. As a result of our homophilic graph formation process, perceptions are most accurate for the poorest which either over- or underestimate actual inequality of  $G = 0.5$  slightly. The Gini coefficient is conventionally defined as the ratio of (unweighted) mean differences in the incomes within the perception set of an individual to twice the mean income within this group. Homophilic graph formation now lets those unweighted mean differences increase linearly at most, while the mean incomes increase exponentially due to the exponential distribution by which incomes are initialised. As a result, the ratio falls almost monotonically. This results is not only plausible due to its accordance with stylised fact (iii) but might also correspond with the empirical evidence on perception

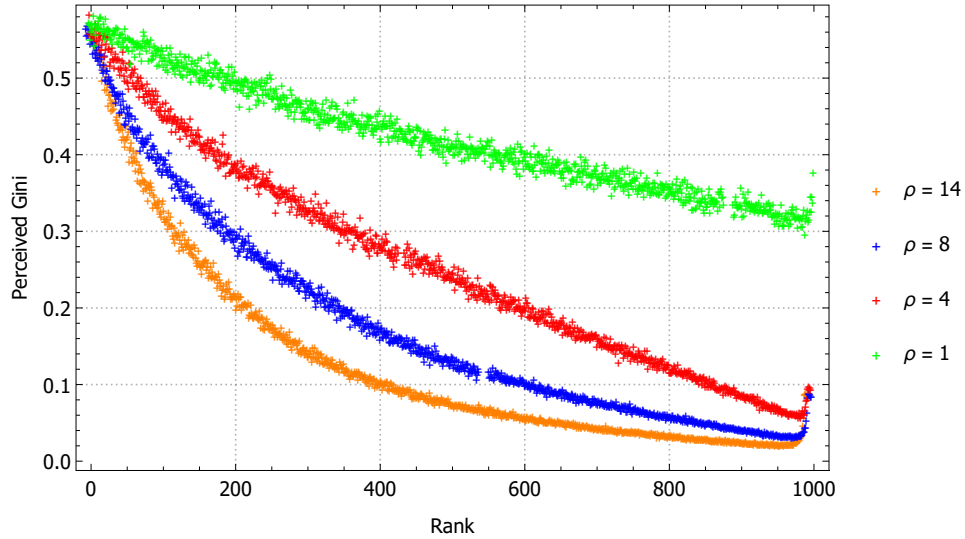


Figure 5.13: Plot of inequality perceptions against the income rank. Almost all individuals underestimate true inequality with a Gini of 0.5. Degrees of underestimation vary, though, where bias increases approximately monotonically in income rank. The intuition for this is that homophilic graph formation lets unweighted inequality (absolute income differences) increase only linearly in income rank, but the reference standard (mean income) increases exponentially.

formation. One of the most prominent hypotheses on perception formation from stimuli is the Weber-Fechner law (Fechner 1862) which indicates that perceived differences in stimuli need to be proportional to the baseline of a given stimulus to be recognisable. The phenomenon is well-established not only for sensory stimuli (Formankiewicz and Mollon 2009; Pienkowski and Hagerman 2009) but also finds use in marketing research on price responses (Sirvanci 1993; Snell et al. 1995). In this framework, one can also understand a decreasing perceived Gini as the change in stimuli (the unweighted differences in incomes of the perception set) do not increase in the same way as the baseline of stimuli (the mean incomes of this perception set) and is thus also consistent with the psychological microevidence.

### 5.4.4 Perceived Global Inequality

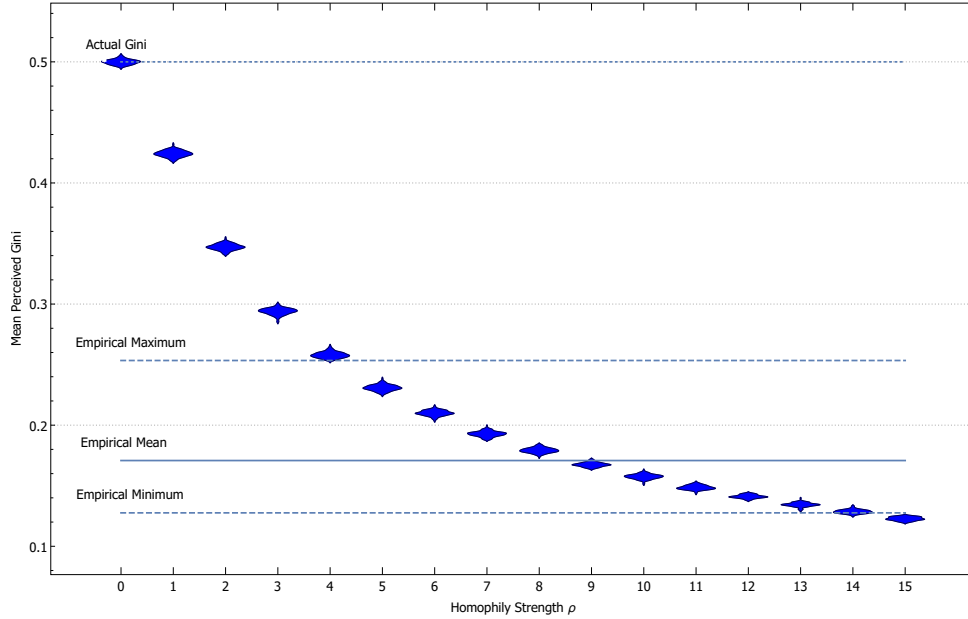


Figure 5.14: The figure shows the violin plots for the cross-sectional average of individual inequality perceptions per Monte Carlo run of our graph model. The actual unbiased Gini of  $G = 0.5$  is indicated by the dotted line. The dashed vertical lines correspond to the empirical sample minimum and maximum, while the bold line corresponds to the sample mean. We find that varying the homophily  $\rho$  parameter can fully quantitatively account for the variation in empirical perceptions.

For further validation, we also examine whether our graph generating process can *quantitatively* replicate empirical perception patterns. We use the mean, minimum and maximum for inequality perceptions on a national level calculated yearly for a large sample of 32 OECD countries in a 30 year time-span by Choi (2019).<sup>12</sup> Over all countries, they find a minimum perceived Gini of  $G_{min} = 0.1276$ , a mean perceived Gini of  $G_{mean} = 0.1708$  and a maximum perceived Gini of  $G_{max} = 0.2534$ . In analogy to their empirical results, we average over the Gini perceptions of all individuals. As we show in Figure 5.14, we find that our process can fully account for their empirical findings and the variation between inequality perceptions by only varying the homophily parameter  $\rho$ . We also note that the sample average of national inequality perceptions implies a homophily degree  $\rho \in [8; 9]$ . Yet, our findings imply considerable cross-country variation in homophily that ranges between  $\rho \approx 4$  to  $\rho \approx 14$ .

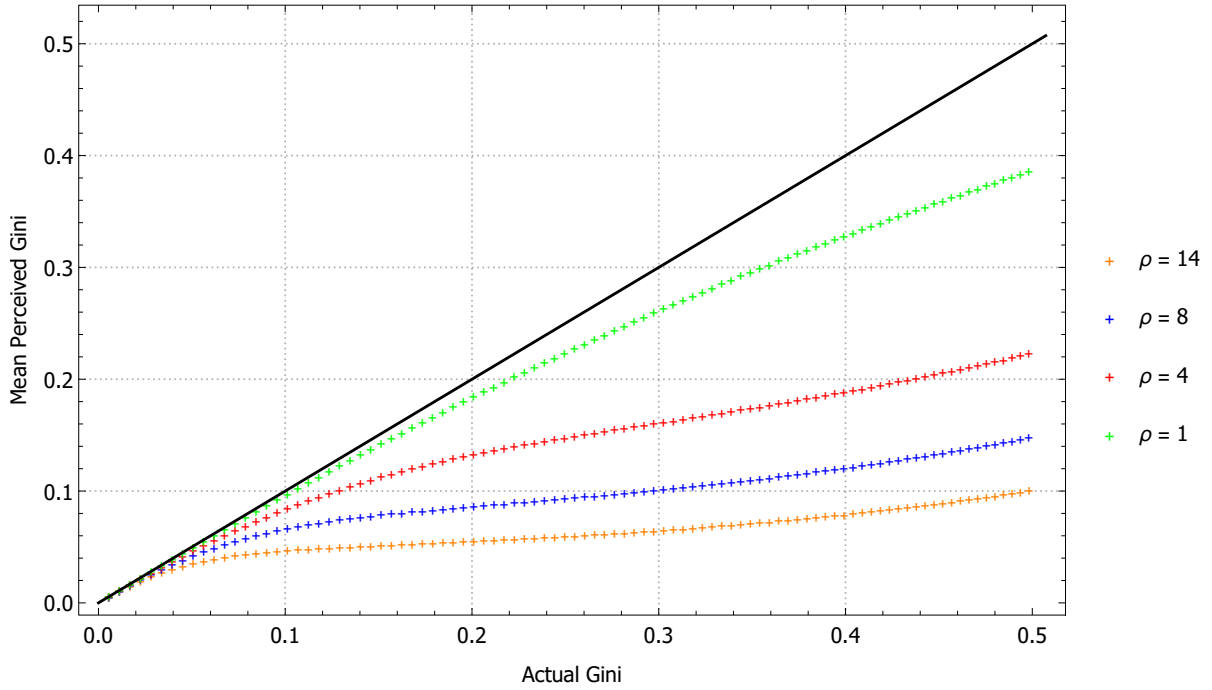


Figure 5.15: The figure plots varying degrees of true inequality against perceived inequality. The solid 45° line corresponds to the true Gini coefficient for direct visual comparison. Apart from extremely low homophily level, increases in actual inequality induce a much lower increase in perceived inequality. For moderate and large degrees of homophily, the schedule reaches a plateau rather quickly, i.e., perceived inequality responds extremely slowly to changes in actual inequality.

#### 5.4.5 Perception Dynamics

To analyse perception dynamics, that is, the reaction of inequality perceptions to changes in actual inequality, we need to initialise the model with another distribution, as the exponential has a fixed Gini of about 0.5, irrespective of its precise parametrisation. We use the log-normal as another benchmark and vary the dispersion parameter  $\sigma$  to simulate changes in the Gini coefficient which is another distribution typically used to describe the skewed nature of empirical income distributions (Knell and Stix 2020). As we see, apart from implausibly low degrees of inequality, changes in actual inequality cause far less than a one-to-one change in perceived inequality. Especially for higher degrees of homophily, we find that the schedule quickly reaches a plateau, where inequality perceptions are now extremely persistent with respect to increases in actual inequality. Our model thus is consistent with stylised fact (iv) as our last test of validity.

The persistence in perceptions occurs because homophily becomes more binding and segregation stronger when actual inequality increases. This mechanism leads *ceteris paribus* to a decrease in perceived inequality which offsets a direct impact of objective inequality

<sup>12</sup>For details and descriptives of their sample, cf. Choi (2019), especially Appendix B2.

on subjective perception. Compare, for illustration, the two regimes close to a completely egalitarian income distribution near  $G = 0$  and relatively high degrees of inequality near  $G = 0.5$ . The egalitarian state is close to a random network, as homophilic segregation presupposes income differences. Small changes in actual inequality are thus not strongly reflected in segregation and almost fully impact perceived inequality, leading to a one-to-one correspondence of perceived and actual inequality in this neighbourhood. For large degrees of actual inequality and large homophily, changes in actual inequality immediately impact segregation, leading to a plateau and very persistent perceptions. Notably, this mechanism is not only consistent with the empirical evidence in terms of its emergent outcome; several recent studies by Reardon and Bischoff (2011); Chen et al. (2012) and Tóth et al. (2021) examine the mechanism directly and show that economic inequality tends to increase (spatial) segregation. A fruitful avenue for further research could be the time-scale on which this channel works, with more laggard segregation responses obviously decreasing the space for inequality-enhancing policies.

#### 5.4.6 Segregation Patterns

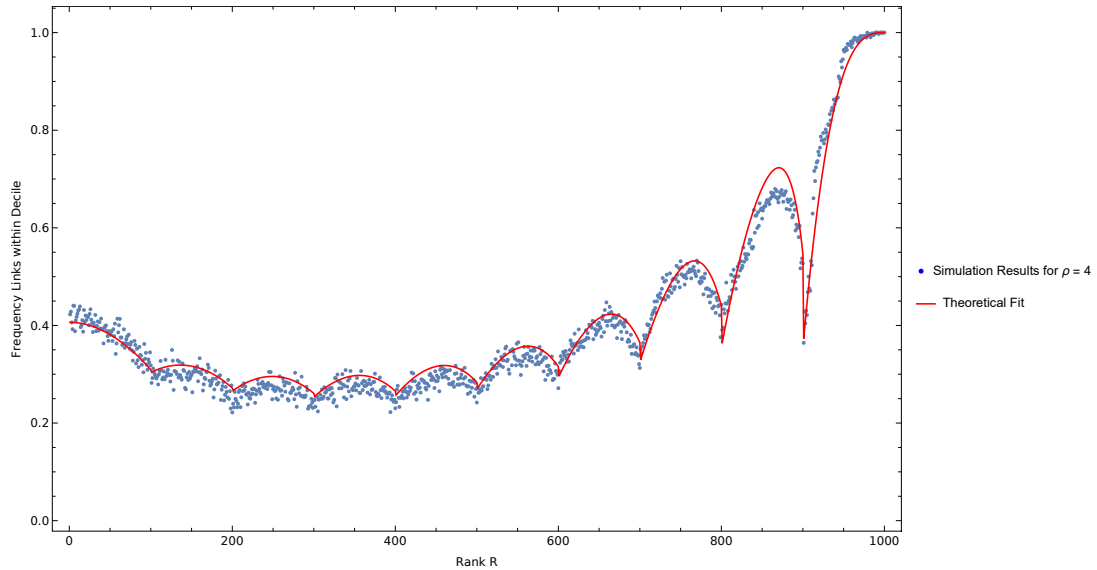


Figure 5.16: The figure plots our segregation measure, measured as the proportion of links of a node to nodes in the same decile, along the rank distribution. The theoretical fit is obtained for the assumption that nodes choose their neighbours themselves without other nodes choosing links incoming to them. The goodness of fit shows that this is indeed a reasonable assumption. Segregation exhibits distinct and non-trivial patterns both regarding global as well as local maxima.

We measure segregation as the proportion of links an individual  $i$  has in their own decile as  $\Delta_i$  as one particular way to measure ‘selectivity’ without access to behavioural linkage parameters. This constitutes a straightforward but standard way to measure segregation



and is easily transformed into normalised measures of segregation like the E-I index  $\Xi_i$  which is defined as the difference between the share of between-group links and the share of within-group links (Bojanowski and Corten 2014).<sup>13</sup> Figure 5.16 plots the simulated segregation statistics as well as a theoretical fit for  $\rho = 4$ . For analytical convenience, the superposed red line plots the probabilities that an individual chooses another agent to link to within their own decile as a first pick, so the total choice set consists of 999 other individuals, and does not account for the possibility that other agents already link to the agent in question, in contrast to our algorithm. 5.A details the derivation. The goodness of fit demonstrates that these incoming-links do not exhibit a significant effect on segregation patterns and tend to average out in the aggregate, showing that our analytical approximation is indeed reasonable.

Segregation exhibits two distinct patterns along the rank distribution. Firstly, we find that segregation exhibits a skewed U-shape and increases, especially for the richest decile, which is almost completely disconnected from the other groups. In this sense, our graph formation process endogenously creates echo chambers for the richest whose information sets do not cover the poorer population at all. This results from the fact that the richest part of the population is extremely selective in choosing their link-neighbours, as we have shown in Section 5.3. Secondly, we also find a rather strong variation together with local maxima within deciles. This finding might be, however, spurious and a partial artefact of boundary effects at decile boundaries. As 5.B shows analytically, individuals will choose those sets of link-neighbours with the highest probability that are distributed symmetrically around them in rank. Thus, individuals exactly at the decile boundary will most likely select a set with half of their neighbours across the boundary. Individuals closer to the centre of a decile, on the other hand, will by the same token choose with highest probability link-neighbours within their own decile. Arbitrarily pre-defined group boundaries can thus create within-group variability in commonly used indices like the E-I index that nevertheless exhibits desirable statistical features at an aggregate level (Bojanowski and Corten 2014). These findings extend well beyond income deciles, as variables like age group, place of living, gender, education or ethnicity are likely strongly correlated with income. Studies using E-I type indices to detect homophily in other variables might hence create spurious results if income homophily is also present. The relevance of such boundary effects has increasingly also been recognised in applied work (Hvidberg et al. 2020). Whenever dimensions have a cardinal scale like income, it might therefore prove more fruitful to use a rolling-window type of estimation, where within-groups are defined in relation to the individual in question, such as a fixed number of income ranks or a fixed income rank interval around theirs.

---

<sup>13</sup>Both measures are simple linear transformations of each other due to  $\Xi_i = 1 - 2\Delta_i$ .

### 5.4.7 Summary of Generating Mechanisms

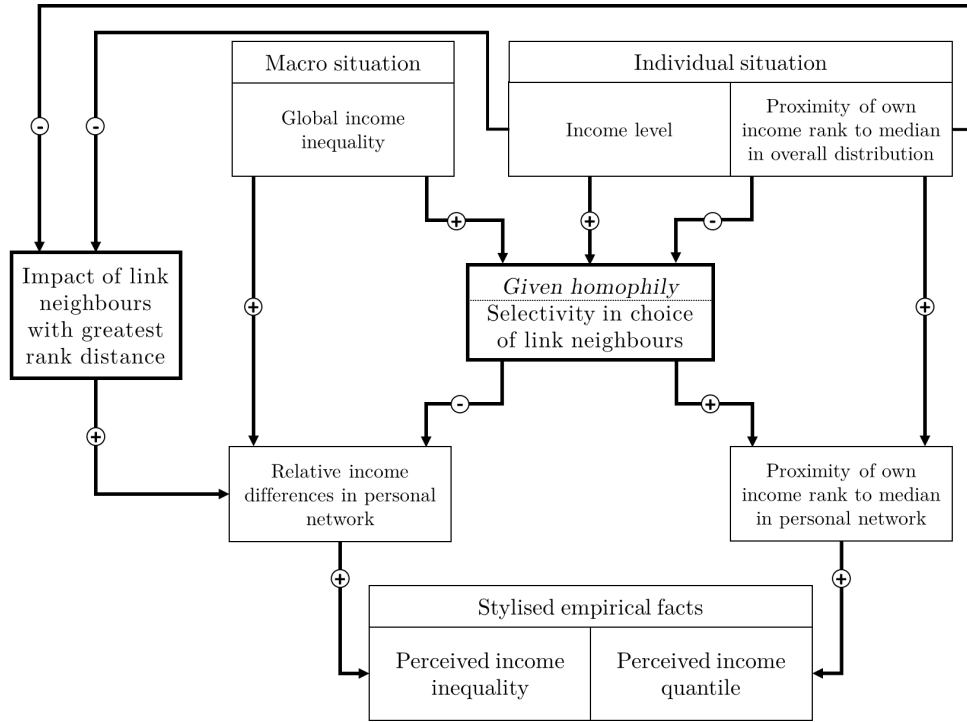


Figure 5.17: Causal pathway depicting the mechanisms of individual selectivity in choice of link-neighbours and subjective perception in the model given a left-skewed income distribution

For non-biased individual samples, one would expect agents to both estimate global inequality correctly in aggregate and also locate their true income quantile. However, homophily triggers link selectivity and hence biased samples which in turn causes inequality perceptions based on an agent's income level and rank (cf. Figure 5.17)

Selectivity in link formation depends on global income inequality. Furthermore, agents whose income is further from the global median income are more selective in their link-neighbours and so are agents with the higher income ranks, as an artefact of the cut-off in the exponential selection function at the low end of the income distribution (cf. the skewed U-shape in Figures 5.3 and 5.4). Such link-neighbour selection generates personal networks for each agent in which this agent tends to have the median income and where income rank differences are relatively small.

Moreover, the extent of relative income differences in one's personal network now depends on the characteristics of the agent in question. Firstly, the impact of link-neighbours with great rank differences to the perceiving agent is larger if this perceiving agent and consequently the majority of link-neighbours have a low income themselves since local inequality calculations weight income differences by the local mean. Secondly, agents

close to the global income distribution median *ceteris paribus* perceive lower inequality levels, for the greatest rank distances tend to be smaller in these cases.

Overall, in a homophilic linking regime, the complex interaction of actual income inequality and of individual absolute income and distribution rank cause biases in income level self-rating and inequality assessment that aggregate to biased underestimation of inequality. Due to the interplay of factors that feed into individual selectivity in choice of link-neighbours, the relation between objective income structure and individual perceptions is non-monotonic and not trivial but requires case-based assessment.

## 5.5 Discussion

Our parsimonious model provides an epistemically possible how-possibly explanation of the stylised empirical facts regarding inequality perceptions that we identified in the literature. Individuals who evaluate their immediate social environment without bias can misperceive their own rank in the overall income distribution as well as global inequality. Homophilic formation of the immediate environment suffices to fully explain the discrepancy between actual and perceived inequality since a rising level of actual inequality causes higher selectivity in link-formation. Moreover, the further away someone's income rank is from the global median and the higher their income, the more selective they are in their choice of link-neighbours.

Thus, public misperceptions are not necessarily driven by limitations in information processing, e.g. a behavioural tendency to place oneself near the median of social hierarchy, but by limited information sets the individuals exhibit for inductive reasoning. The seemingly subtle distinction between constraints on information processing and information sets carries important policy implications: When it is the limits of available information and not limits in cognitive ability driving misperceptions, informational treatments may be successful, as also the empirical literature suggests (Cruces et al. 2013; McCall et al. 2017; Karadja et al. 2017). Such treatments can either consist of delivering information about income inequality itself or facilitating the formation of more diverse contacts in order to overcome the segregation by income that our model finds. In other words, this means breaking up echo chambers that are caused by humans drawing confidence in their beliefs only from repeated observations while ignoring a potential lack of diversity in sources (Foster et al. 2012; Schwarz et al. 2016).

Educating individual citizens about their information deficit and providing ways of overcoming it is important from a democracy theory perspective. For example, (Rawls 2005, p. 224) requires “presently accepted general beliefs” as basis for arguments in the public forum. However, while one can assess the income inequality objectively without any room for disagreement if using all globally available information, citizens who work only with their individual information will agree on a belief about the Gini that underestimates its

actual value. Hence, the lack of individual information access inhibits deliberation about the level inequality and its changes, e.g., in response to past policy measures.

For the first time, we are able to infer the composition of these reference groups from readily available observational data on perceived inequality to inform both empirical investigations as well as more comprehensive model-building in other regards. Directly investigating perception networks might provide a possible remedy for the problem that identification of interaction effects within survey data in the form of time series is hindered by sample sizes that are typically one or two orders of magnitude too low to distinguish noise from true interaction (Alfarano and Milaković 2012). Our main empirical prediction is that the homogeneity of social groups, the fraction of links to agents within the own income decile, roughly follows a U-shaped pattern with a massive decrease in diversity for the richest and (a quantitatively much more modest one) for the poorest agents. This finding might point to an endogenously emerging ‘elite discourse’ with almost no transmission of information to the poorer 90% of the population.

Our cross-country analysis shows that there exists considerable variability in implied homophily levels. There are several possible candidates to explain this variability, such as cultural norms, diversity in media and political representation or spatial segregation. Regarding spatial segregation patterns, Thorstein Veblen made the farsighted observation as early as 1899 that urbanisation should increase diversity in social contacts, since cities are the place “where the human contact of the individual is widest and the mobility of the population is greatest” (Veblen 1899, p. 66). Thus, perceived inequality should *ceteris paribus* be higher in urban areas resulting from the higher average income diversity per perception network, a testable hypothesis and thus a possible avenue for further research in spatial economics. This is also what the rather scarce existing evidence for Vietnam and Central and Eastern Europe suggests (Mahajan et al. 2014; Binelli and Loveless 2016). In this way, spatial and perception network segregation might therefore overlap and interact (Newman et al. 2018; Tóth et al. 2021), and policy affecting the allocation of land could thus also exhibit unintentional effects on perceptions. We leave detailed analyses on these determinants for further research.

In terms of theory, our empirically validated random geometric graphs might provide an ideal microfoundation for theories of consumption as dependent on the relative income position and for which shocks that affect local income compositions lead to ‘expenditure cascades’ (Duesenberry 1949; Veblen 1899; Frank et al. 2014). This new mechanism might shed light on the disputed link between economic inequality and growth. We will address these questions in further research.

Finally, our model presents a way of generating random geometric graphs defining both the distribution of the feature that determines linking probability between any pair of nodes and a minimum degree value for each node. Put differently, we apply a Barabási-Albert (1999) Preferential-Attachment type procedure, which is intuitive for social scientists, to

features other than degree and get network graphs that can be analysed using readily available methods from random geometric graph theory. To the best of our knowledge, there is no such way yet.

While it is reassuring that a scale-transformation to a log-scale delivers the same salient ‘middle-class bias’ as our baseline specification, log-transformations in general strike us as a very parsimonious way to capture scale-dependence in choice. This is not only relevant for inequality perceptions but for essentially all variables and features where perception of stimuli is plausibly dependent on scale. In general, it is possible to apply the generating procedure to features other than income that exhibit different distributional patterns. This approach will hopefully inspire future studies of expectation formation, e.g., regarding inflation or business sentiment. In these fields, identification of the relevant perception networks might be a crucial step to bring macroeconomic theory currently mostly building on atomistic rational expectations and empirical studies, that find little support for those types of expectations, closer together (Pesaran and Weale 2006). Hence, we also provide a toolkit for analysing the impact of homophily regarding any specified feature on network generation (and potential interactions on the resulting network) given a particular distribution of this feature that is well-grounded in the behavioural economics literature. In sum, we anticipate our theoretical framework to generate numerous avenues for further studies both theoretically, regarding graph- and expectation formation, as well as empirically, pertaining to the determinants of homophily and possible policy measures, to information transmission and to the effects of inequality on aggregate consumption.

## 5.A Linkage Probabilities for Homophilic Networks

**Heuristic Derivation.** Linkage Probabilities.

Consider an arbitrary node  $K_i$  indexed  $i \in 0, \dots, N-1$  as their rank  $R$  increasing in income that is part of a graph  $G$  with  $N$  nodes characterised by adjacency matrix  $A$ . Let  $I_i$  denote their income, where  $f_\lambda(\cdot)$  defines the PDF of a exponential probability distribution defined over the real half-line  $(0, \infty)$  with parameter  $\lambda > 0$  and  $F_\lambda(\cdot)$  the corresponding CDF. The quantile function for any population share  $p$  and with parameter  $\lambda > 0$  for an exponential distribution is given as

$$\phi_\lambda(p) = \frac{-\log[1-p]}{\lambda}. \quad (5.6)$$

We assume without loss of generality that  $\lambda = 1$  for normalisation. Calculated  $\rho$  values therefore need to be scaled up by the inverse of the mean income,  $\lambda^{-1}$  for empirical application.

The quantile of a node with income  $I_i$  can be approximated by their rank  $R$ , such that  $p \approx R/N$  as a discrete approximation of the continuous probability density which holds for large  $N$ . We want to derive the probability that a node  $i$  with rank  $R \in \mathbb{N}_0^+$  connects to a node  $j$  with a distance of  $d$  to node  $i$ . Expressing the weights as defined in Section 5.3 now in the form of quantiles, we get

$$w_{ij} = \exp[-\rho |\phi_{\lambda=1}(R/N) - \phi_{\lambda=1}((R+d)/N)|] \quad (5.7)$$

$$= \exp[-\rho |\log(1 - (R/N)) - \log(1 - (R+d)/N)|]. \quad (5.8)$$

Assume first that  $d > 0$ , that is,  $j$  is richer than  $i$ . Simplifying the weights yields for  $d > 0$ :

$$w_{ij} = \exp[-\rho (\log(1 - (R/N)) - \log(1 - (R+d)/N))] \quad (5.9)$$

$$= \exp[\log \left( \left( \frac{N-R}{N} \right)^{-\rho} \right) - \log \left( \left( \frac{N-R-d}{N} \right)^{-\rho} \right)] \quad (5.10)$$

$$= \left( \frac{N-R}{N-R-d} \right)^{-\rho} \quad (5.11)$$

$$= \left( \frac{N-R-d}{N-R} \right)^{\rho} \quad (5.12)$$

Analogously, we get for  $d < 0$

$$w_{ij} = \exp[-\rho (\log(1 - (R+d)/N) - \log(1 - (R/N)))] \quad (5.13)$$

$$= \exp[\log \left( \left( \frac{N-R-d}{N} \right)^{-\rho} \right) - \log \left( \left( \frac{N-R}{N} \right)^{-\rho} \right)] \quad (5.14)$$

$$= \left( \frac{N-R-d}{N-R} \right)^{-\rho} \quad (5.15)$$

$$= \left( \frac{N-R}{N-R-d} \right)^{\rho} \quad (5.16)$$

To translate  $w_{ij}$  into probabilities, we need to normalise by all weights. Note that this is still a (close) approximation of the probabilities of link-formation of a given node  $i$ . Nodes draw their  $C$  link-neighbours from the set of all neighbours. This implies that draws are not independent, as we assume here. Since  $C \ll N$ , however, the effect is marginal. The approximation for the probability below, however, seems to perform quite well which we verify in our subsection on segregation. The probability  $p$  that  $i$  chooses  $j$  as a link-partner can therefore be approximated as

$$p_{ij}(N, R, d) \approx \begin{cases} \left( \frac{N-R}{N-R-d} \right)^{\rho} / \left( \sum_{\tilde{d}=-R+1}^{-1} \left( \frac{N-R}{N-R-\tilde{d}} \right)^{\rho} + \sum_{\tilde{d}=1}^{N-R-1} \left( \frac{N-R-\tilde{d}}{N-R} \right)^{\rho} \right) & \text{for } d < 0, \\ \left( \frac{N-R-d}{N-R} \right)^{\rho} / \left( \sum_{\tilde{d}=-R+1}^{-1} \left( \frac{N-R}{N-R-\tilde{d}} \right)^{\rho} + \sum_{\tilde{d}=1}^{N-R-1} \left( \frac{N-R-\tilde{d}}{N-R} \right)^{\rho} \right) & \text{for } d > 0. \end{cases}$$

Notice that the function behaves as expected and is monotonically decreasing in  $|d| \in \mathbb{N}^+$ . The strength of selection also increases monotonically in the homophily parameter  $\rho$ . For  $\rho = 0$ , we recover the equiprobable case without any decay. The precise functional form of the decay for  $\rho \in \mathbb{R}^+$  is, however, far from trivial and changes along the rank distribution. The right tail of the correspondent density is always a power transformation of a linear function, whereas the left tail for any given  $R$  is a power transformation of a function with hyperbolic decay. In this sense, all nodes are more 'selective' regarding individuals that are poorer than regarding the richer part of the population. To see this, compare the decay for the minimum and the maximum of the distribution for  $\rho = 1$  as a special case. For  $R = 0$ ,  $p_{ij} \propto 1 - (|d|/N)$  with linear decay in  $|d|$ , as there exists only a right tail, while for  $R = N - 1$ ,  $p_{ij} \propto 1/(1 + |d|)$  which decays extremely fast in  $|d|$  by a power function, as there exists only a left tail here. In this sense, the richest individual is far more 'selective' in choosing their (poorer) link-neighbours than the poorest individual choosing their (richer) ones.

The theoretical expected segregation index we compare against our simulation results can be straightforwardly computed from those probabilities. Let  $\delta_i$  be the set of nodes that are in the same group as node  $i$  such as an income quantile. The probability to connect with a link-neighbour  $\tilde{p}_i$  can then again be approximated as

$$\tilde{p}_i(N, R, d) \approx \sum_{j \in \delta_i} p_{ij}. \quad (5.17)$$

## 5.B Perceived Quantiles in Perception Networks

**Proof Sketch.** Pure Homophily implies a Tendency to the Median in Perceived Quantiles.

Consider an arbitrary node  $K_i$  indexed  $i \in 0, \dots, N - 1$  in a graph  $G$  characterised by adjacency matrix  $A$ . Let  $I_i$  denote their income, where  $f_\lambda(\cdot)$  defines the PDF of an exponential probability distribution defined over the real half-line  $(0, \infty)$  with parameter  $\lambda > 0$  and  $F_\lambda(\cdot)$  the corresponding CDF. Let  $M$  be the number of links of node  $K_i$  with  $M$  even. This leaves us with  $\binom{N-1}{M} = S$  possible permutations of link-neighbours. Assume further for  $F_\lambda(I_i)$  that it is between  $1/2 \cdot M/N$  and  $1 - (1/2 \cdot M/N)$ , such that

$$\frac{1}{2} \frac{M}{N} < F_\lambda(I_i) < 1 - \frac{1}{2} \frac{M}{N}. \quad (5.18)$$

Let now  $\theta_{ij}$  be an arbitrary realisation of a permitted set of incomes of nodes to which  $K_i$  links, indexed by  $j$  out of the set of permitted sets  $\Theta_i$  with  $\Theta_i = \{\theta_{i1}, \dots, \theta_{iS}\}$  and size  $S$ .

Assume further that all incomes in  $\theta_{ij}$  are distinct. If link formation is independent of  $I_i$  as the sole characteristic differentiating  $K_i$  from all other nodes, all sets  $\theta_{ij}$  of the same size  $M$  are equally likely with probability  $1/S$  by extension, since  $K_i$  connects to any other node with equal probability. This would be the case for both standard preferential attachment models as well as ER random graphs.

In our model, the probability  $p_{ik}$  that  $A_{ik} = 1$  depends negatively on the absolute distance  $|I_i - I_k|$ , such that  $\partial p_{ik} / \partial |I_i - I_k| < 0$ . By linearity, the probability  $p_{ij}$  of node  $i$  to have  $\theta_{ij}$  as their chosen set of incomes to which she is linked decreases in the sum of absolute differences, that is,  $\partial p_{ij} / \partial \sum_{I_k \in \theta_{ij}} |I_i - I_k| < 0$ . It follows, that  $p_{ij}$  as a local probability of a set of a given length being chosen by homophilic preferential attachment is maximised for a minimisation of  $\sum_{I_k \in \theta_{ij}} |I_i - I_k|$ . Since the benchmark without homophily is equal probability of  $1/S$  for all sets of a given size  $M$ , this condition also maximises the global probability that this set is chosen for a given size  $M$ . Formally, the minimisation problem chooses a set or sets  $\theta_{ij}$  such that

$$\arg \min_{\theta_{ij}} \sum_{I_k \in \theta_{ij}} |I_i - I_k|. \quad (5.19)$$

It remains to be shown that this minimisation leads to the choice of a set  $\theta_{ij}$  for which  $I_i$  is the median value. Let  $\phi_\lambda(\cdot) = \phi(\cdot)$  to simplify notation. The median requires the same number of values above or below  $I_i$  in  $\theta_{ij}$ . With  $M$  links for node  $K_i$  of income rank  $R$  and  $M$  even, this requires  $M/2$  values above and below  $I_i$ . For  $I_i$  as the median being minimising for the absolute distances, this requires i) that there exists no node with rank  $R + M/2 + 1$  such that their income distance to  $K_i$  is less than the income distance from node  $K_i$  to the node ranked  $R - M/2$ . If i) is violated, the node with rank  $R + M/2 + 1$  is part of the distance-minimising set and thus,  $I_i$  is not the median of  $\theta_{ij}$ . The symmetrical condition ii) requires that there is no node with rank  $R - M/2 - 1$  such that its distance to  $K_i$  is less than the distance of  $K_i$  to the node with rank  $R + M/2$ . In terms of a quantile function, we require

$$\phi\left(\frac{R + M/2 + 1}{N}\right) - \phi\left(\frac{R}{N}\right) > \phi\left(\frac{R}{N}\right) - \phi\left(\frac{R - M/2}{N}\right) \quad (5.20)$$

and

$$\phi\left(\frac{R}{N}\right) - \phi\left(\frac{R - M/2 - 1}{N}\right) > \phi\left(\frac{R + M/2}{N}\right) - \phi\left(\frac{R}{N}\right). \quad (5.21)$$



Rearranging yields

$$\phi\left(\frac{R + M/2 + 1}{N}\right) + \phi\left(\frac{R - M/2}{N}\right) > 2\phi\left(\frac{R}{N}\right) > \phi\left(\frac{R + M/2}{N}\right) + \phi\left(\frac{R - M/2 - 1}{N}\right). \quad (5.22)$$

Expressing the left-hand side of inequalities for a generic distribution in (5.22) for an continuous exponential such that  $R/N \approx p$  and substituting the quantile function, we require

$$\frac{-\log[1 - (R + M/2 + 1)/N]}{\lambda} + \frac{-\log[1 - (R - M/2)/N]}{\lambda} > 2 \frac{-\log[1 - (R/N)]}{\lambda} \quad (5.23)$$

The condition  $R/N \approx p$  presupposes  $N$  to be sufficiently large for the discrete realisations of the sample to approximate the quantiles of the continuous exponential distribution. We find this condition fulfilled for several numerical experiments. It is easy to see that the left-hand side condition in (5.22) is fulfilled for a quantile function whose first derivative is monotonically increasing which is the case for  $d\phi_\lambda(p)/dp = 1/((1 - p)\lambda)$  for  $p \in [0, 1)$  and  $\lambda > 0$ . We can also show this by manipulation of (5.23) as

$$\log[(1 - (R + M/2 + 1)/N) \cdot (1 - (R - M/2)/N)] < \log[(1 - (R/N))^2] \quad (5.24)$$

$$\Leftrightarrow (1 - (R + M/2 + 1)/N) \cdot (1 - (R - M/2)/N) < (1 - (R/N))^2 \quad (5.25)$$

which implies

$$1 - \frac{(R + M/2 + 1)}{N} - \frac{R - M/2}{N} + \frac{(R + M/2 + 1)(R - M/2)}{N^2} - 1 + 2\frac{R}{N} - \frac{R^2}{N^2} < 0 \quad (5.26)$$

$$\Leftrightarrow -\frac{1}{N} + \frac{R - M/2 - (M/2)^2}{N^2} < 0 \quad (5.27)$$

$$\Leftrightarrow \frac{R - M/2 - (M/2)^2 - N}{N^2} < 0. \quad (5.28)$$

Since  $R \leq N$  per definition, condition (5.28) is trivially fulfilled. Notice that this implies for an exponential initial distribution,  $I_i$  cannot be below the median in the most likely set. The right hand-side of inequalities (5.22) is a bit more demanding. Stating the condition in terms of the quantile function for an exponential, we get

$$2 \frac{-\log[1 - (R/N)]}{\lambda} > \frac{-\log[1 - (R - M/2 - 1)/N]}{\lambda} + \frac{-\log[1 - (R + M/2)/N]}{\lambda}. \quad (5.29)$$

Simplifying yields

$$(1 - (R/N))^2 < (1 - (R - M/2 - 1)/N) \cdot (1 - (R + M/2)/N) \quad (5.30)$$

$$\Leftrightarrow (1 - (R/N))^2 < \frac{(N - R + M/2 + 1)(N - R - M/2)}{N^2} \quad (5.31)$$

$$\Leftrightarrow (N - R)^2 < N^2 - RN - NM/2 - RN + R^2 + RM/2 + NM/2 - RM/2 - M^2/4 + N - R - M/2 \quad (5.32)$$

$$\Leftrightarrow 0 < N - R - M/2 - M^2/4 \quad (5.33)$$

$$\Leftrightarrow \frac{R}{N} < 1 - \frac{1}{2} \frac{M}{N} - \frac{M^2}{4N}. \quad (5.34)$$

For our discrete sample,  $F_\lambda(I_i) \approx R/N$  which reveals that condition (5.34) is only a slightly more demanding condition than boundary condition (5.18) that guarantees the possibility of  $I_i$  being a median in the first place and only differs by  $M^2/4N$ . Since we typically assume  $M \ll N$ , this term vanishes. Indeed, for a realistic baseline scenario with  $N = 1,000$  and  $M = 4$ , the condition is fulfilled for the poorest 99% of the population and thus for the vast majority. Together with the lower boundary condition (5.18), the tendency to place themselves in the middle should exist for about 98% of the population and thus the vast majority. Minimising absolute deviations for an exponential income distribution and  $M \ll N$  thus entails choosing sets that let  $I_i$  be the median of  $\theta_{ij} \cup I_i$  for almost all  $I_i$ . While the strength of this mechanism will of course be dependent on  $\partial p_{ij} / \partial \sum_{I_k \in \theta_{ij}} |I_i - I_k|$ , the median is the most likely outcome for any homophilic network as the perceived quantile for the vast majority of nodes.

## 5.C Linkage Based on Logarithmised Income Differences

In this appendix, we discuss the case where agents seek to minimise relative rather than absolute income differences in tie formation, i.e., weights for tie-formation are inversely proportional to the absolute distance in log income. Agent  $j$ 's weight in agent  $i$ 's draw  $\tilde{w}_{ij}$  is thus determined by

$$\tilde{w}_{ij} = \frac{1}{\exp[\rho |\ln[I_j] - \ln[I_i]|]}. \quad (5.35)$$

The scale-transformation by logarithmising tends to offset the property of the exponential income distribution to exhibit much higher (absolute) income differences in its upper tail than in the lower parts of the distribution, since the natural log has negative second derivative. For moderate to high levels of  $\rho$  agents will have a strong tendency to choose an equal number of agents above or below them in income that is roughly homogeneous across the income distribution (apart from agents located at the boundary). For self-perceptions, this implies a ‘middle-class bias’ per stylised facts (i) and, by extension, (ii), as is also readily verified by simulation with the results below in Figures 5.18 to 5.21.

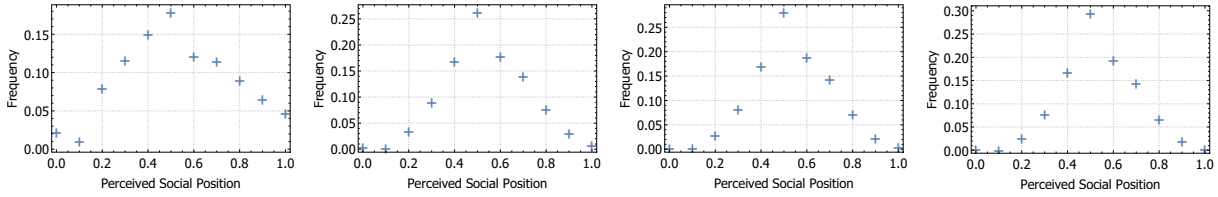


Figure 5.18: Density of Perceived Quantiles for  $\rho = 1$ . Figure 5.19: Density of Perceived Quantiles for  $\rho = 4$ . Figure 5.20: Density of Perceived Quantiles for  $\rho = 8$ . Figure 5.21: Density of Perceived Quantiles for  $\rho = 14$ .

*Note:* The Figures report the perceived social positions for  $\rho \in \{1; 4; 8; 14\}$  with 10 bins each and for a choice function with log incomes in its argument. All Figures exhibit significant deviation from the benchmark with equal frequencies. While the tendency is apparent in all panels, its strength expectedly increases and perceptions are more narrowly clustered around the middle categories the higher  $\rho$  is.

Logarithmising income in the weight function as in eq. (5.35) bears enormous consequences for inequality rather than self-perceptions, though. This is what Figure 5.22 below shows. Apart from agents at the respective upper and lower boundaries, the approximately homogeneous segregation tendency across the income distribution manifests itself in roughly homogeneous inequality perceptions for all considered  $\rho \geq 4$ . For low  $\rho$ , the behaviour of local Ginis is thus broadly consistent with one aspect of stylised fact (iii), namely that perceived inequality tends to decrease in income rank, as is also shown in Figure 5.22 for  $\rho = 1$ . However, poor agents then drastically overestimate inequality in violation of the second aspect of stylised fact (iii) that (almost) all agents somewhat drastically underestimate inequality.

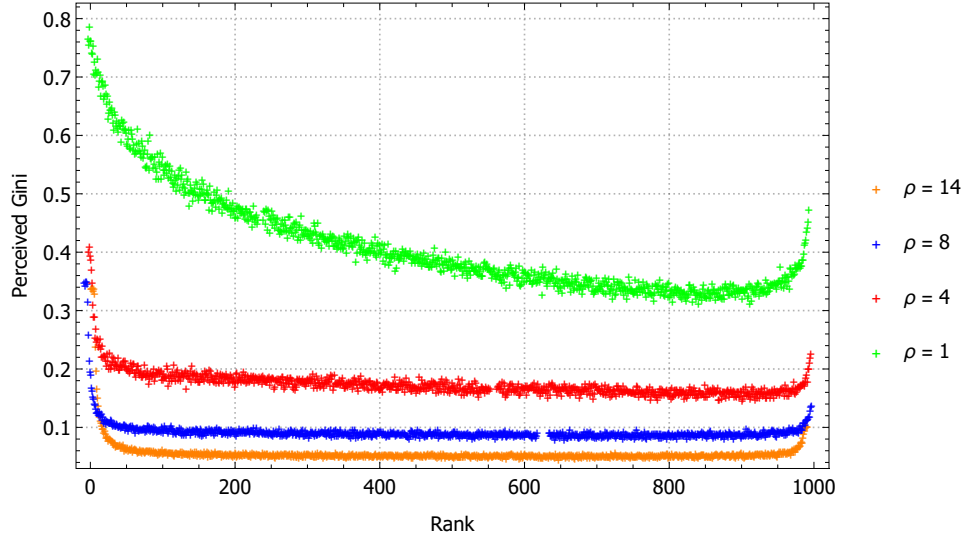


Figure 5.22: Plot of inequality perceptions against the income rank. Almost all individuals underestimate true inequality with a Gini of 0.5 in all cases of  $\rho > 0$ . For moderate to high homophily strengths  $\rho$  between 4 and 14, there is little covariation of income rank and perceptions (except for the inflated perceptions at the upper and lower boundary).

Heuristically, the behaviour of the local Gini coefficients is a direct consequence of the behaviour of an exponential income distribution on a log-scale. Incomes increase locally linear in rank around the mean income (of  $\lambda = 1$ ) but increase much faster superlinearly near the upper and lower tails of the distribution, as Figure 5.23 shows.

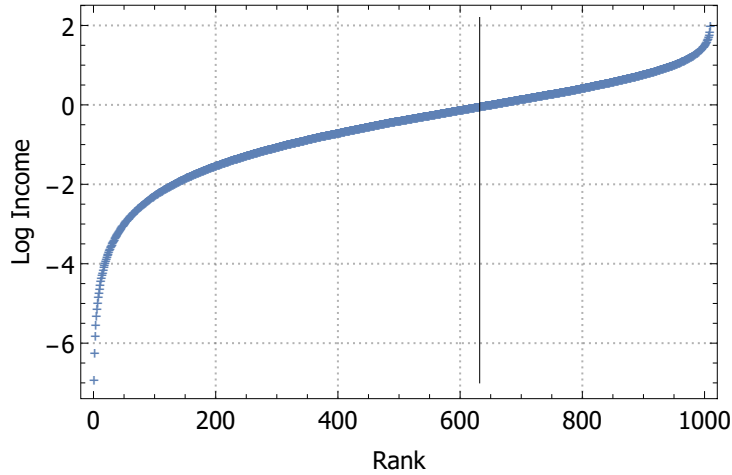


Figure 5.23: Exponential income distribution in logs as a function of income rank. Around the mean income of  $\lambda = 1$ , indicated by the vertical gridline, income grows approximately linearly in rank, while growth is superlinear at the tails.

This shape of the distribution of logarithmised incomes implies that the agents with income at the lower tail will have a relatively strong probability of observing incomes around the mean, where income ranks do not change inclusion probabilities very much.

Since these values are, from the perspective of low-income agents, extreme values, this higher probability of observing incomes within this locally linear region translates into higher expected inequality perceptions, as is evident from the boundary effects in Figure 5.22. A symmetric argument applies to the upper tail of the distribution, although with a more attenuated effect, since the mean income is closer to the maximum value in income rank due to the skewness of the exponential. Hence, the values around the mean are not as extreme from the perspective of the agents with maximal incomes, leading to a less pronounced increase in upper boundary inequality perceptions. Agents with log incomes located in the locally linear region exhibit the strongest tendency to perceive incomes close to them in rank and exhibit the most pronounced dislike of extreme perceptions at the upper and lower tail. In this sense, the bias against extreme values increases with decreasing distance to the mean income which indicates that inequality perceptions decrease until the mean income is reached and increase afterwards. Increasing  $\rho$  then disproportionately affects extreme distances due to the exponential nature of the weights, mitigating any differences in perceptions caused by them. For  $\rho \geq 4$ , any differential behaviour is then barely visible in local inequality perceptions within Figure 5.22, in contrast to both stylised fact (iii) and our results for choice with absolute income differences.

## References

- Alfarano, S. and Milaković, M. (2012). Identification of interaction effects in survey expectations: a cautionary note. *Studies in Nonlinear Dynamics & Econometrics*, 16(4).
- Alicke, M. D., Zell, E., and Bloom, D. L. (2010). Mere categorization and the frog-pond effect. *Psychological Science*, 21(2):174–177.
- Anufriev, M., Chernulich, A., and Tuinstra, J. (2018). A laboratory experiment on the heuristic switching model. *Journal of Economic Dynamics and Control*, 91:21–42.
- Anufriev, M. and Hommes, C. (2012). Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics*, 4(4):35–64.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.
- Barde, S. and Van Der Hoog, S. (2017). An empirical validation protocol for large-scale agent-based models. *Bielefeld Working Papers in Economics and Management*, 04-2017.
- Bartels, L. M. (2018). *Unequal democracy: The political economy of the new gilded age*. Princeton University Press.

- Binelli, C. and Loveless, M. (2016). The urban–rural divide: Perceptions of income and social inequality in Central and Eastern Europe. *Economics of Transition*, 24(2):211–231.
- Boguná, M., Pastor-Satorras, R., Díaz-Guilera, A., and Arenas, A. (2004). Models of social networks based on social distance attachment. *Physical Review E*, 70(5):056122.
- Bojanowski, M. and Corten, R. (2014). Measuring segregation in social networks. *Social Networks*, 39(1):14–32.
- Borgatti, S. P. and Foster, P. C. (2003). The network paradigm in organizational research: A review and typology. *Journal of Management*, 29(6):991–1013.
- Boucher, V. (2015). Structural homophily. *International Economic Review*, 56(1):235–264.
- Buckingham, J. T. and Alicke, M. D. (2002). The influence of individual versus aggregate social comparison and the presence of others on self-evaluations. *Journal of Personality and Social Psychology*, 83(5):1117.
- Campbell, D. E. (2013). Social networks and political participation. *Annual Review of Political Science*, 16:33–48.
- Caplan, B. (2011). *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton University Press.
- Cepić, D. and Tonković, Ž. (2020). How social ties transcend class boundaries? Network variability as tool for exploring occupational homophily. *Social Networks*, 62:33–42.
- Chater, N., Zhu, J.-Q., Spicer, J., Sundh, J., León-Villagrà, P., and Sanborn, A. (2020). Probabilistic biases meet the Bayesian brain. *Current Directions in Psychological Science*, 29(5):506–512.
- Chen, W.-H., Myles, J., and Picot, G. (2012). Why have poorer neighbourhoods stagnated economically while the richer have flourished? Neighbourhood income inequality in Canadian cities. *Urban Studies*, 49(4):877–896.
- Chiang, Y.-S. (2011). Judgment of distributional inequality in networks. *Social Networks*, 33(4):342–349.
- Chiang, Y.-S. (2015a). Good samaritans in networks: An experiment on how networks influence egalitarian sharing and the evolution of inequality. *PLOS ONE*, 10(6):e0128777.
- Chiang, Y.-S. (2015b). Inequality measures perform differently in global and local assessments: An exploratory computational experiment. *Physica A: Statistical Mechanics and its Applications*, 437:1–11.
- Choi, G. (2019). Revisiting the redistribution hypothesis with perceived inequality and redistributive preferences. *European Journal of Political Economy*, 58:220–244.
- Clark, A. E. and Senik, C. (2010). Who compares to whom? The anatomy of income comparisons in europe. *Economic Journal*, 120(544):573–594.

- Cohen, J. (1979). Socio-economic status and high-school friendship choice: Elmtown's youth revisited. *Social Networks*, 2(1):65–74.
- Collier, P. and Tuckett, D. (2020). Narratives as a coordinating device for reversing regional disequilibrium. *Rebuilding Macroeconomics Working Paper Series*, 22.
- Canyon, M. J. and Muldoon, M. R. (2006). The small world of corporate boards. *Journal of Business Finance & Accounting*, 33(9-10):1321–1343.
- Cruces, G., Perez-Truglia, R., and Tetaz, M. (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, 98:100–112.
- Dall, J. and Christensen, M. (2002). Random geometric graphs. *Physical Review E*, 66(1):016121.
- Davis, G. F., Yoo, M., and Baker, W. E. (2003). The small world of the American corporate elite, 1982–2001. *Strategic Organization*, 1(3):301–326.
- Dawtry, R. J., Sutton, R. M., and Sibley, C. G. (2015). Why wealthier people think people are wealthier, and why it matters: From social sampling to attitudes to redistribution. *Psychological Science*, 26(9):1389–1400.
- De Paula, A. (2017). Econometrics of network models. In *Advances in Economics and Econometrics: Theory and Applications: Eleventh World Congress*, volume 1, pages 268–323. Cambridge University Press.
- Drăgulescu, A. and Yakovenko, V. M. (2001). Evidence for the exponential distribution of income in the USA. *European Physical Journal B*, 20(4):585–589.
- Duesenberry, J. S. (1949). *Income, saving, and the theory of consumer behavior*. Harvard University Press.
- Engelhardt, C. and Wagener, A. (2018). What do Germans think and know about income inequality? A survey experiment. *Socio-Economic Review*, 16(4):743–767.
- Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, 4(5):41–60.
- Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5(1):17–60.
- Evans, M. D. and Kelley, J. (2004). Subjective social location: Data from 21 nations. *International Journal of Public Opinion Research*, 16(1):3–38.
- Fagiolo, G., Guerini, M., Lamperti, F., Moneta, A., and Roventini, A. (2019). Validation of agent-based models in economics and finance. In Beisbart, C. and Saam, N. J., editors, *Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives*, pages 763–787. Springer International Publishing, Cham.
- Fechner, G. T. (1862). *Elemente der Psychophysik*. Breitkopf und Härtel, Leipzig.

- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2):117–140.
- Flieth, B. and Foster, J. (2002). Interactive expectations. *Journal of Evolutionary Economics*, 12(4):375–395.
- Formankiewicz, M. A. and Mollon, J. (2009). The psychophysics of detecting binocular discrepancies of luminance. *Vision research*, 49(15):1929–1938.
- Foster, J. L., Huthwaite, T., Yesberg, J. A., Garry, M., and Loftus, E. F. (2012). Repetition, not number of sources, increases both susceptibility to misinformation and confidence in the accuracy of eyewitnesses. *Acta psychologica*, 139(2):320–326.
- Frank, R. H., Levine, A. S., and Dijk, O. (2014). Expenditure cascades. *Review of Behavioral Economics*, 1(1–2):55–73.
- Franke, R. and Westerhoff, F. (2017). Taking stock: A rigorous modelling of animal spirits in macroeconomics. *Journal of Economic Surveys*, 31(5):1152–1182.
- Fronczak, A., Fronczak, P., and Hołyst, J. A. (2004). Average path length in random networks. *Physical Review E*, 70(5):056110.
- Gabaix, X. (2020). A Behavioral New Keynesian Model. *American Economic Review*, 110(8):2271–2327.
- Galaskiewicz, J. (2007). Has a network theory of organizational behaviour lived up to its promises? *Management and Organization Review*, 3(1):1–18.
- Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.
- Gilbert, N. and Troitzsch, K. (2005). *Simulation for the social scientist*. McGraw-Hill Education (UK).
- Gimpelson, V. and Treisman, D. (2018). Misperceiving inequality. *Economics and Politics*, 30(1):27–54.
- Golub, B. and Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *Quarterly Journal of Economics*, 127(3):1287–1338.
- Gräbner, C. (2018). How to relate models to reality? An epistemological framework for the validation and verification of computational models. *Journal of Artificial Societies and Social Simulation*, 21(3).
- Grüne-Yanoff, T. and Verreault-Julien, P. (2021). How-possibly explanations in economics: anything goes? *Journal of Economic Methodology*, 28(1):114–123.
- Hall, P. A. and Soskice, D. (2001). *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford University Press.
- Hamilton, M. J., Milne, B. T., Walker, R. S., Burger, O., and Brown, J. H. (2007). The complex structure of hunter–gatherer social networks. *Proceedings of the Royal Society B: Biological Sciences*, 274(1622):2195–2203.



- Huckfeldt, R. R. (1983). Social contexts, social networks, and urban neighborhoods: Environmental constraints on friendship choice. *American Journal of Sociology*, 89(3):651–669.
- Humphries, M. D. and Gurney, K. (2008). Network ‘small-world-ness’: a quantitative method for determining canonical network equivalence. *PLOS ONE*, 3(4):e0002051.
- Hvidberg, K. B., Kreiner, C., and Stantcheva, S. (2020). Social position and fairness views. *NBER Working Paper Series*, 28099.
- Iacono, R. and Ranaldi, M. (2021). The nexus between perceptions of inequality and preferences for redistribution. *Journal of Economic Inequality*, 19(1):97–114.
- ISSP Research Group (2016). International Social Survey Programme (ISSP) 1985–2014. *GESIS Data Archive, Cologne*.
- Kalmijn, M. (1991). Status homogamy in the United States. *American Journal of Sociology*, 97(2):496–523.
- Kalmijn, M. and Flap, H. (2001). Assortative meeting and mating: Unintended consequences of organized settings for partner choices. *Social Forces*, 79(4):1289–1312.
- Karadja, M., Mollerstrom, J., and Seim, D. (2017). Richer (and holier) than thou? The effect of relative income improvements on demand for redistribution. *Review of Economics and Statistics*, 99(2):201–212.
- Karimi, F., Génois, M., Wagner, C., Singer, P., and Strohmaier, M. (2018). Homophily influences ranking of minorities in social networks. *Scientific reports*, 8(1):1–12.
- Kelley, J. and Evans, M. (1995). Class and Class Conflict in Six Western Nations. *American Sociological Review*, 60(2):157–178.
- Kenworthy, L. and McCall, L. (2008). Inequality, public opinion and redistribution. *Socio-Economic Review*, 6(1):35–68.
- Kim, H., Huh, S., Choi, S., and Lee, Y. (2018). Perceptions of inequality and attitudes towards redistribution in four East Asian welfare states. *International Journal of Social Welfare*, 27(1):28–39.
- Kleinberg, J. (2001). Small-world phenomena and the dynamics of information. *Advances in neural information processing systems*, 14:431–438.
- Knell, M. and Stix, H. (2020). Perceptions of inequality. *European Journal of Political Economy*, 65:101927.
- Kogut, B. and Walker, G. (2001). The small world of Germany and the durability of national networks. *American Sociological Review*, 66(3):317–335.
- Kraus, M. W., Rucker, J. M., and Richeson, J. A. (2017). Americans misperceive racial economic equality. *Proceedings of the National Academy of Sciences*, 114(39):10324–10331.
- Kuhn, A. (2019). The subversive nature of inequality: Subjective inequality perceptions and attitudes to social inequality. *European Journal of Political Economy*, 59:331–344.

- Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., and Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *Journal of Politics*, 62(3):790–816.
- Lazarsfeld, P. F. and Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. In Berger, M., Abel, T., and Charles, H., editors, *Freedom and control in modern society*, pages 18–66. D. Van Nostrand Company.
- Lewis, K., Gonzalez, M., and Kaufman, J. (2012). Social selection and peer influence in an online social network. *Proceedings of the National Academy of Sciences*, 109(1):68–72.
- Lodge, M., Steenbergen, M. R., and Brau, S. (1995). The responsive voter: Campaign information and the dynamics of candidate evaluation. *American Political Science Review*, 89(2):309–326.
- Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, 15(3):215–233.
- Lux, T. (2009). Rational forecasts or social opinion dynamics? Identification of interaction effects in a business climate survey. *Journal of Economic Behavior & Organization*, 72(2):638–655.
- MacCarron, P., Kaski, K., and Dunbar, R. (2016). Calling Dunbar’s numbers. *Social Networks*, 47:151 – 155.
- Mahajan, S., Dinh, V. T., Rab, H., Trieu, V. Q., Nguyen, L. V., Kwakwa, V., and Verghis, M. (2014). Taking stock: an update on Vietnam’s recent economic developments. Technical report, Washington, D.C.: World Bank Group.
- Mäki, U. (2009). MISSing the world. Models as isolations and credible surrogate systems. *Erkenntnis*, 70:29–43.
- Malacarne, T. (2017). Rich friends, poor friends: Inter-socioeconomic status friendships in secondary school. *Socius*, 3:1 – 13.
- Manski, C. F. and McFadden, D. (1981). *Structural analysis of discrete data with econometric applications*. MIT Press Cambridge, MA.
- Mayer, A. and Puller, S. L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics*, 92(1-2):329–347.
- McCall, L., Burk, D., Laperrière, M., and Richeson, J. A. (2017). Exposure to rising inequality shapes americans’ opportunity beliefs and policy support. *Proceedings of the National Academy of Sciences*, 114(36):9593–9598.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.
- Moore, C. and Newman, M. E. (2000). Epidemics and percolation in small-world networks. *Physical Review E*, 61(5):5678.
- Newman, B. J., Shah, S., and Lauterbach, E. (2018). Who sees an hourglass? Assessing citizens’ perception of local economic inequality. *Research and Politics*, 5(3).

- Newman, M. E. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409.
- Newman, M. E. (2009). Random graphs with clustering. *Physical Review Letters*, 103(058701).
- Newman, M. E., Strogatz, S. H., and Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(026118).
- Osberg, L. and Smeeding, T. (2006). "Fair" inequality? Attitudes toward pay differentials: The United States in comparative perspective. *American Sociological Review*, 71(3):450–473.
- Page, B. I. and Shapiro, R. Y. (1993). The rational public and democracy. In Marcus, G. E. and Hanson, R., editors, *Reconsidering the Democratic Public*. Pennsylvania State University Press.
- Pesaran, M. H. and Weale, M. (2006). Survey expectations. In Elliott, G., Granger, C. W., and Timmermann, A. G., editors, *Handbook of Economic Forecasting*, volume 1, pages 715–776. Elsevier.
- Pienkowski, M. and Hagerman, B. (2009). Auditory intensity discrimination as a function of level-rove and tone duration in normal-hearing and impaired subjects: The "mid-level hump" revisited. *Hearing research*, 253(1-2):107–115.
- Rawls, J. (2005). *Political liberalism*. Columbia University Press.
- Reardon, S. F. and Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, 116(4):1092–1153.
- Reutlinger, A., Hangleiter, D., and Hartmann, S. (2018). Understanding (with) toy models. *British Journal for the Philosophy of Science*, 69(4):1069–1099.
- Sanborn, A. N. and Chater, N. (2016). Bayesian brains without probabilities. *Trends in Cognitive Sciences*, 20(12):883–893.
- Schwarz, N., Newman, E., and Leach, W. (2016). Making the truth stick & the myths fade: Lessons from cognitive psychology. *Behavioral Science & Policy*, 2(1):85–95.
- Silva, A. C. and Yakovenko, V. M. (2004). Temporal evolution of the "thermal" and "superthermal" income classes in the USA during 1983–2001. *Europhysics Letters (EPL)*, 69(2):304.
- Sirvanci, M. B. (1993). An empirical study of price thresholds and price sensitivity. *Journal of Applied Business Research (JABR)*, 9(2):43–49.
- Snell, J., Gibbs, B. J., and Varey, C. (1995). Intuitive hedonics: Consumer beliefs about the dynamics of liking. *Journal of Consumer Psychology*, 4(1):33–60.
- Snijders, T. A. (2011). Statistical models for social networks. *Annual Review of Sociology*, 37:131–153.
- Stevenson, R. T. and Duch, R. (2013). The meaning and use of subjective perceptions in studies of economic voting. *Electoral Studies*, 32(2):305–320.

- Sugden, R. (2009). Credible worlds, capacities and mechanisms. *Erkenntnis*, 70(1):3–27.
- Talaga, S. and Nowak, A. (2020). Homophily as a process generating social networks: Insights from social distance attachment model. *Journal of Artificial Societies and Social Simulation*, 23(2):6.
- Tao, Y., Wu, X., Zhou, T., Yan, W., Huang, Y., Yu, H., Mondal, B., and Yakovenko, V. M. (2019). Exponential structure of income inequality: evidence from 67 countries. *Journal of Economic Interaction and Coordination*, 14(2):345–376.
- Thal, A. (2017). Class isolation and affluent Americans’ perception of social conditions. *Political Behavior*, 39(2):401–424.
- Törnqvist, L., Vartia, P., and Vartia, Y. O. (1985). How should relative changes be measured? *The American Statistician*, 39(1):43–46.
- Tóth, G., Wachs, J., Di Clemente, R., Jakobi, Á., Ságvári, B., Kertész, J., and Lengyel, B. (2021). Inequality is rising where social network segregation interacts with urban topology. *Nature Communications*, 12(1):1–9.
- Uzzi, B. and Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology*, 111(2):447–504.
- Veblen, T. (1899). *The Theory of the Leisure Class*. Republished (2001), New York: The Modern Library.
- Watts, D. J. (1999). Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, 105(2):493–527.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442.
- Weeden, K. A. and Cornwell, B. (2020). The small-world network of college classes: implications for epidemic spread on a university campus. *Sociological Science*, 7:222–241.
- Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge University Press.
- Zell, E. and Alicke, M. D. (2009). Contextual neglect, self-evaluation, and the frog-pond effect. *Journal of Personality and Social Psychology*, 97(3):467.
- Zhou, W.-X., Sornette, D., Hill, R. A., and Dunbar, R. I. (2005). Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B: Biological Sciences*, 272(1561):439–444.

# Perception and Privilege

---

Joint work with D. M. Mayerhoffer.

Published under Creative Commons Attribution 4.0 License:

Mayerhoffer, D. M. and Schulz, J. (2022). Perception and privilege. *Applied Network Science*, 7(1):1–25. <https://doi.org/10.1007/s41109-022-00467-x>

This paper benefited greatly from discussions at the Complex Networks and their Applications 2021 conference and the peer-review of an earlier version of this manuscript from two anonymous reviewers for this conference. We are also grateful for helpful comments at the NetPLACE seminar series, the Networks in a Global World 2022 conference, the 14th Annual Conference of the International Network for Analytical Sociology and the 26th Forum for Macroeconomics and Macroeconomic Policies, especially Elisa Belotti and Jan Fuhse as well as Anna Gebhard and Bettina Gregg. We also thank Carsten Källner for his able research assistance. All remaining errors are, of course, our own.

---

## Abstract

Inequality perceptions differ along racial and gendered lines. To explain these disparities, we propose an agent-based model of localised perceptions of the gender and racial wage gap in networks. We show that the combination of homophilic graph formation and estimation based on locally limited knowledge can replicate both the underestimation of the gender or racial wage gap that empirical studies find and the well-documented fact that the underprivileged perceive the wage gap to be higher on average and with less bias. Similarly, we demonstrate that the underprivileged perceive overall inequality to be higher on average. In contrast to this qualitative replication, we also show that the effect of homophilic graph formation is quantitatively too strong to account for the empirically observed effect sizes within a recent Israeli sample on perceived gender wage gaps. As a parsimonious extension, we let agents estimate using a composite signal based on local and global information. Our calibration suggests that women place much more weight on the (correct) global signal than men, in line with psychological evidence that people adversely affected by group-based inequities pay more attention to global information about the issue. Our findings suggest that (educational) interventions about the global state of gender equality are much more likely to succeed than information treatments about overall inequality and that these interventions should target the privileged.

---

---

## Contents of this Paper

6.1	Introduction . . . . .	167
6.2	Related Literature . . . . .	168
6.3	Model . . . . .	170
6.3.1	Agent Initialisation . . . . .	171
6.3.2	Network Generation . . . . .	172
6.3.3	Individual Perception . . . . .	174
6.3.4	Validation Strategy . . . . .	176
6.4	Results . . . . .	177
6.4.1	Mean Local Perceptions . . . . .	177
6.4.2	Disaggregating Local Perceptions . . . . .	180
6.4.3	Empirical Calibration and Composite Signal . . . . .	182
6.4.4	The Impact of Privilege on General Inequality Perceptions . . . . .	185
6.5	Discussion . . . . .	186
6.A	Appendix: Calibrating the Composite Wage Gap . . . . .	189
	References . . . . .	191

---

## 6.1 Introduction

Perceptions matter. There is now widespread agreement that it is *perceived* inequities, not necessarily actual ones behind redistributive preferences and policy decisions (Hauser and Norton 2017; Choi 2019). Apart from biased perceptions of empirical wealth and income distributions, the literature increasingly also documents misperceptions of the gender and racial inequities (Kraus et al. 2017; Malul 2021), relating, in particular, to wage gaps between groups of the privileged and underprivileged. This finding carries a vital policy implication: If the general public overestimates the achieved progress in both racial and gender equality, the issue might never surface in public debate (Settele 2022). There also exists considerable theoretical and empirical evidence that individual perceptions of unfairness trigger wage negotiations and, consequently, wage growth (Akerlof and Yellen 1990; Pfeifer and Stephan 2019). Misperceptions about the wage gap might thus contribute to its persistence documented in Akchurin and Lee (2013) - and addressing this persistence would require addressing its biased perceptions.

Despite their relevance for both market outcomes and policy, there is no formal model of belief formation in this regard to the best of our knowledge. There are two potential channels that may (jointly or independently) lead to the biased individual perceptions: firstly, biased information processing of a correct signal, i.e., the true global (population-wide) wage gap; secondly, incorrect individual signals that come from a skewed sample. Such a sample might fail to accurately represent the population, e.g., due to social networks segregated by income Kraus et al. (2017).

We propose a simple network model based on homophilic attachment initially employed for perceptions of overall inequality by Schulz et al. (2022) and find that the second channel alone is sufficient to explain an underestimation of the wage gap, even given fully rational, unbiased individual information processing. In fact, model agents who rely only on the biased information from their individual sample underestimate the wage gap more severely than empirical studies find. Thus, we combine the biased local signal with the correct wage gap as global signal.

By calibrating our model with data from a recent sample in Malul (2021), we find that both the local and the global channels matter but that their relative importance differs between the privileged and underprivileged. Our theoretical findings imply that informational treatments appear to have a much more significant effect when targeted on the privileged. By contrast, our results suggest that purely local perceptions are sufficient to generate the empirically observed underestimation of overall inequality. Here, we find that the underprivileged unambiguously perceive higher inequality and are more accurate on average in their assessment. This insight offers a novel explanation for why women and ethnic minorities consistently favour more redistribution. Rather than assuming that they have different value orientations or want the welfare state to redistribute more out of self-

interest, we argue that they simply have a clearer picture of the problem and perceive inequality to be higher.

The remainder of this paper is organised as follows: Section 6.2 discusses related empirical findings that our model aims to replicate. We introduce the model in section 6.3. Section 6.4 introduces our basic results both on the mean perceptions for the privileged and underprivileged class as well as the highly non-monotonic relationship between income, neighbourhood disassortativity and perceptions, as well as our calibration exercise with Israeli data from Malul (2021). Section 6.5 spotlights policy implications of our results and discusses avenues for further theoretical and empirical work.

## 6.2 Related Literature

This paper considers perceptions of the wage gap between the privileged and underprivileged as emerging from endogenously evolving reference groups. The yet relatively scarce literature on such perceived wage gaps has mainly focused on gender and racial wage gaps. For the gender wage gap, the extant literature suggests that the population appears to underestimate the actual extent of gender-based inequality (Malul 2021) but that this effect is primarily driven by male bias, with women having much higher and more accurate perceptions on average (Hampton and Heywood 1993; Foley et al. 2002; García-González et al. 2019; Malul 2021). In line with these findings, women exhibit a stronger tendency to perceive both the overall level of inequality (Andreoni and Vesterlund 2001; Alesina and Giuliano 2011; Bavetta et al. 2019) and gender-based inequalities (Pfeifer and Stephan 2019) to be unjustified.<sup>1</sup> Regarding perception formation about the gender wage gap, the literature suggests two predominant channels: Individuals form their perceptions from lived experience in (potentially) gender-diverse situations (Auspurg et al. 2017) but are also generally aware of the global wage gap, e.g., through national media (Williams et al. 2010; Furnham and Wilson 2011).

The findings on perceptions of racial inequality are largely in line with those for the gender wage gap. The general population underestimates the extent of racial inequality across various dimensions, including the wage gap (Kraus et al. 2017, 2019; Alesina et al. 2021; Davidai and Walker 2021). These studies also tend to find that Black people who are adversely affected by the wage gap tend to perceive higher wage gaps and are, generally, more accurate in its assessment (Kraus et al. 2017, 2019; Davidai and Walker 2021). Kraus et al. (2017) also find that the richest White people tend to underestimate the racial gap most strongly. Intriguingly, they speculate and provide some evidence that this might be related to the diversity in social networks that is particularly low for

---

<sup>1</sup>Note that the effect of gender on perceptions of fairness reverses in Pfeifer and Stephan (2019) when controls for the hourly wage are introduced. Since our model-output consists of the unconditional effects without controlling for income, the unconditional effect in Pfeifer and Stephan (2019) is the relevant benchmark, though.



the wealthiest part of the population. Carter et al. (2019) corroborate this hypothesis and show that diversity in students' friendship networks empirically exhibits a positive association with perceptions of racial injustice. Apart from this local channel, the efficacy of informational treatments reporting aggregates also suggests the relevance of global averages on individual perceptions (Haaland and Roth 2021).

Homophily in social contacts may be the product of chance or a by-product of individual behaviour with different aims - or individuals may be consciously homophilic in their tie formation, e.g., to reduce strategic uncertainty (Kets and Sandroni 2019). Empirically, there is explicit income homophily surfacing in, e.g. mobile phone communication (Xu et al. 2021; Leo et al. 2016; Fixman et al. 2016). Moreover, willing and accidental homophily in other dimensions can be represented as income homophily because income proximity is a proxy for proximity in other dimensions: Empirical studies show factors like residential area (Hu and Liang 2022; Harting and Radi 2020), education (Smith et al. 2014; Leo et al. 2016), lifestyle (Virtanen et al. 2007), or even health (Krieger 1992) to correlate with income; and there is obviously also a correlation between income and ethnicity (Chandra 2000), i.e. one possible dimension to observe privilege on with regard to our model. Of course, the named factors are also correlated with each other and we do not aim to argue for a potential direction of causality. Instead, we utilise the implication of the empirically well established correlation: Even if people do not willingly choose their social contacts based on income proximity, their actual choices amount to social contacts as if selected homophilic in income.

To validate our model that we introduce in section 6.3, we thus consider two stylised facts extracted from the empirical literature: (i) The general population appears to underestimate the extent of gender or racial inequality and (ii) this underestimation is mainly driven by the bias of the privileged agents. As for generating mechanisms, both local knowledge in the form of lived experience and global signals, e.g. through the national media, appear to be necessary.

While perceptions thus appear to be detached from actual gender and racial inequities, we still need to consider the actual levels of inequality as a benchmark. Gender and racial wage gaps were generally between 10 % and 40% in the last decade for industrialised countries which is also the range we will consider (Kunze 2018; Stelzner and Bahn 2022). Moreover, it is now a well-established stylised fact that empirical wage distributions follow an exponential law which has been demonstrated in numerous studies for a vast range of countries and sample periods, i.e., for Australia (Banerjee et al. 2006), Romania (Derzsy et al. 2012; Oancea et al. 2017), the European Union as a whole (Jagielski and Kutner 2013), Japan (Nirei and Souma 2007), the US (Nirei and Souma 2007; Dragulescu and Yakovenko 2001a,b; Banerjee and Yakovenko 2010; Dos Santos 2017; Schneider 2015; Shaikh 2017; Shaikh and Jacobo 2020; Silva and Yakovenko 2004) as well as the UK (Dragulescu and Yakovenko 2001b; Banerjee and Yakovenko 2010; Dos Santos 2017; Schneider

2015; Shaikh 2017; Shaikh and Jacobo 2020; Silva and Yakovenko 2004). Most notably, Tao et al. (2019) confirm for their whole sample of 67 countries that the exponential law provides an excellent fit for empirical wage distributions.<sup>2</sup> For our purposes, though, not only the overall distribution is relevant but also the wage distribution of privileged and underprivileged wage-earners considered separately. In a foundational contribution, Shaikh et al. (2014) indeed also demonstrated that the wage distribution of men and women, as well as the ones of Whites, Hispanics and African-American considered separately, are well approximated by this exponential distribution for US data. While the wage distributions of the underprivileged class thus have lower average income, they empirically follow the same functional form as the wage distributions of the privileged and the population as a whole.

After an in-depth validation of our model, we directly apply the general model structure to a long-standing policy issue: Attitudes towards redistribution. A large literature now documents that women (Linos and West 2003; Quadagno and Blekesaune 2003; Cusack et al. 2006; Alesina and Giuliano 2011) and ethnic minorities (Alesina and Giuliano 2011; Kinder and Winter 2001; Morgan and Kelly 2017) are generally much more supportive of welfare state measures and redistributive policies. The main hypotheses to explain these gender and racial differentials are based on different value orientations or on the fact that the precarious working and living conditions by women and ethnic minorities let those groups support redistributive measures out of self-interest (Quadagno and Blekesaune 2003). Similarly, Linos and West (2003) suggest that women tend to support welfare state measures because they are more likely to work within the public sector, therefore directly benefitting from the expansion of the welfare state. We provide a complementary explanation: Our model unambiguously implies that the underprivileged perceive higher levels of inequality and are less biased in their perception. Thus, they might favour redistribution simply because they have a much clearer picture of the existing inequalities and encounter them more frequently or saliently in their daily lives.

## 6.3 Model

The present agent-based computational model aims to study the impact of social tie formation that is homophilic in wage on individual perceptions of a wage gap and wage inequality in general. The link formation is important because model agents use their ego network as a sample to form their perceptions. The model and analyses here extend Mayerhoffer and Schulz (2022) who in turn build on a linking procedure introduced by Schulz et al. (2022) to study perceptions of inequality in general. It comprises of three distinct

---

<sup>2</sup>The upper tail of 1 to 5 % empirical income distributions appears to follow a power-law or Pareto distribution (Silva and Yakovenko 2004). The income concentrated there is primarily *capital income*, though, while we are concerned with *wages* and labour income. Henceforth, we exclude this small minority from our consideration.

phases that take place once per simulation run and in sequential order: (I) Agent initialisation and group-specific income allocation, (II) network generation through agents' homophilic linkage, and (III) individual wage gap perception and network evaluation, which comprises of (III.a) comparing wages between groups within one's neighbourhood and (III.b) processing of the correct population-wide wage gap as a global signal. Figure 6.1 gives an overview of these phases to guide the content-oriented descriptions below; for further technical details, consider the NetLogo implementation of the simulation model and the output data.

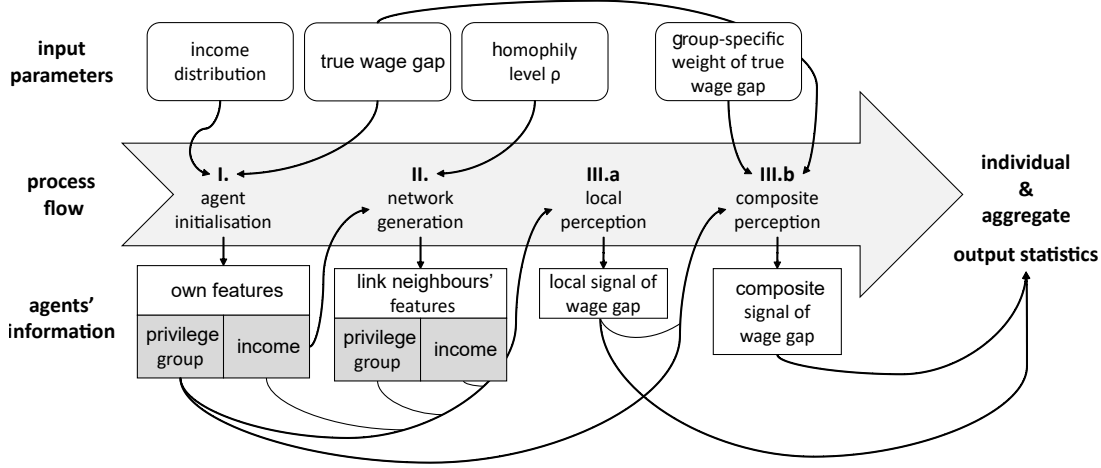


Figure 6.1: Overview over the simulation processes

### 6.3.1 Agent Initialisation

The simulated population consists of 1,000 agents in the model that belong to either the privileged group/class (e.g., White people, men) or the underprivileged one (e.g., BiPoC, women). At initialisation, each agent draws their income from an exponential distribution with a mean of  $\lambda = 1$ . This pre wage gap income distribution is identical for all agents regardless of their privilege group. As discussed in section 6.2, such a distribution normalises the empirically observed (pre-tax or market) wage distributions in various industrialised countries for the vast majority of individuals (Tao et al. 2019). We use the empirical finding that the wage distribution of the total population as well as of the relevant population subgroups is exponential for our model initialisation. The wage distribution we use for initialisation can thus be considered as constituting a representative sample taken from empirical wage distributions. The distribution of underprivileged earners is downscaled by a wage gap  $g \in [0, 1]$ , in line with the empirical findings of Shaikh et al. (2014). Privileged agents thus always retain the income they drew while the underprivileged agents' income is reduced by a proportion  $g \in [0, 1]$ . The findings on the general population's income distribution (Tao et al. 2019) as well as our simulation results ensure that for any wage gap, the unconditional distribution of incomes irrespective of

privilege also follows an exponential law. They allow us to construct a general model that applies equally to perceived gender and racial wage gaps based on privilege classes, since decomposition of the income distribution along both these dimensions delivers the same exponential distributional regularity.

### 6.3.2 Network Generation

Each agent selects five other agents as link-neighbours. Therefore, the model does not impose links on agents according to a global rule, but, like in real-world networks, they *create* their links themselves. If an agent  $i$  picks agent  $j$  who had themselves picked  $i$  before that, the already existing link between the two agents remains untouched, but  $i$  does not pick another neighbour instead of  $j$ . Consequently, each agent has at least five link-neighbours (i.e., social contacts) but may have more. This number is empirically validated, as we intend to represent the closest layer of intense contacts (Zhou et al. 2005; Hamilton et al. 2007; Mac Carron et al. 2016)<sup>3</sup>. Moreover, we have carried out sensitivity analyses and found a higher number of links to have little impact on overall simulation results. Since this closest layer may consist of differently related people for different individuals (Karlsson et al. 2005), a link in our model can represent any social relationship: It simply indicates that an agent knows their link neighbour’s income. Because links are undirected and equally weighted within the model, this knowledge about income is always mutual. When selecting link neighbours, an agent does not care about the potential selectee’s group belonging. However, agents do care about the potential drawee’s income, as link-generation is homophilic in income. Homophily generally describes the tendency to link to similar individuals. This mechanism is important for information diffusion Larson (2017). Moreover Milli (2021) employs homophily in an agent-based model as influence factor of perception. Income constitutes one particularly relevant dimension McPherson et al. (2001) of homophily and the only dimension we consider here. Namely, agent  $j$ ’s weight in agent  $i$ ’s draw is denoted by  $\Omega_{ij}$  and determined as follows:

$$\Omega_{ij} = \frac{1}{\exp[\rho |Y_j - Y_i|]} \quad (6.1)$$

The relative weights in the draws are a function of the homophily strength and the respective income levels:  $Y \in \mathbb{R}^+$  denotes the income of an agent.  $\rho \in \mathbb{R}_0^+$  denotes the homophily strength in income selection, externally set, and identical for all agents.  $\rho = 0$  represents a random graph, and for an increasing value of  $\rho$ , an agent becomes ever more likely to pick link-neighbours with incomes closer to their own. Put differently, for each agent, there is a likelihood ranking over all other agents based on income differences. Which agent occupies which rank and especially to which privilege class they belong

---

<sup>3</sup>For a recent review on the extensive literature on ‘Dunbar’s number’, cf. the first section of Mac Carron et al. (2016).

depends on  $Y$  and thus on the wage gap  $g$ .  $\rho$  determines how strongly  $\Omega$  reacts to those income differences.

The link function's exponential character ensures that those with the largest income differences become unlikely picks even at low to moderate homophily strengths. The choice of an exponential weighting function might seem arbitrary, but upon closer inspection, we find that translated into the probability of  $i$  choosing  $j$ , this weighting is equivalent to the discrete choice approach developed and popularised by Manski (1981). The homophily parameter  $\rho \in [0, \infty)$  is then simply the intensity of the choice parameter. To translate weights into probabilities, we normalise by all weights for all agents, i.e.,

$$p_{ij} = \frac{\exp[-\rho \cdot |Y_j - Y_i|]}{\sum_{k \in M \setminus i} \exp[-\rho \cdot |Y_k - Y_i|]}, \quad (6.2)$$

with  $M \setminus i$  as the set of all agents except  $i$  with size  $N - 1$ .<sup>4</sup> Considering probabilities rather than weights in (6.2) is intuitively interpretable, with  $\rho = 0$  implying equiprobable picks with  $p_{ij} = 1/(N - 1), \forall j \in M \setminus i$ , and thus indeed a random graph, while  $\rho \rightarrow \infty$  implies that  $p$  approaches unity for  $j$  with minimum income distance to the income of  $i$  and 0 for all other agents. Manski (1981) show that the discrete choice rule follows directly from random utility theory, i.e., agents maximising utility and utility functions being decomposable into an observable and unobservable component and both being uncorrelated. In our case, the observable component the agents minimise would be the income differences, with the unobservable part being all the attributes from which our agent in question would benefit due to their social connection. It follows that the linkage function in eq. (6.1) is plausibly microfounded in a utility-maximising framework.<sup>5</sup> In this sense, our weighting function is a tractable reduced-form representation of the empirically established homophily. Falk and Knell (2004) demonstrate that endogenous reference standards like ours can also emerge from an optimisation framework with utility functions that incorporate plausible relative motives.

Figure 6.2 illustrates the linkage probabilities implied by the weighted draw based on the exponentially distributed income levels. The local maxima of individual linkage probability densities exhibit a bi-modal shape with peaks at the highest and lowest rank but are also heavily skewed to the left, i.e., agents with high incomes are most selective in their link picks. We understand selectivity as the tendency to select agents that are close in income rank as link-neighbours. For an extensive analysis of this linkage behaviour, see Schulz et al. (2022). General selectivity increases with  $\rho$ , and for  $g > 0$ , selectivity in

---

<sup>4</sup>Note that this is, strictly speaking, only the probability of the first choice of agent  $i$ , since we consider drawing without replacement. In particular, the weighting function does not account for the possibility that other agents already link to the agent in question, in contrast to our algorithm. Since the number of agents is rather large, the effect appears to cancel out in the aggregate, as is verified in (Schulz et al. 2022).

<sup>5</sup>The discrete choice approach has now become the workhorse formalisation within behavioural macroeconomics. Cf. e.g. Franke and Westerhoff (2017) for a recent survey.

income may also mean selectivity in groups. Most notably, this implies that our model generates gender-based and racial homophily as a (from a modelling perspective desirable) byproduct of income homophily since income differences are partially determined by privilege class. Indeed, empirically, homophily operates along many different dimensions, including race and gender (McPherson et al. 2001).

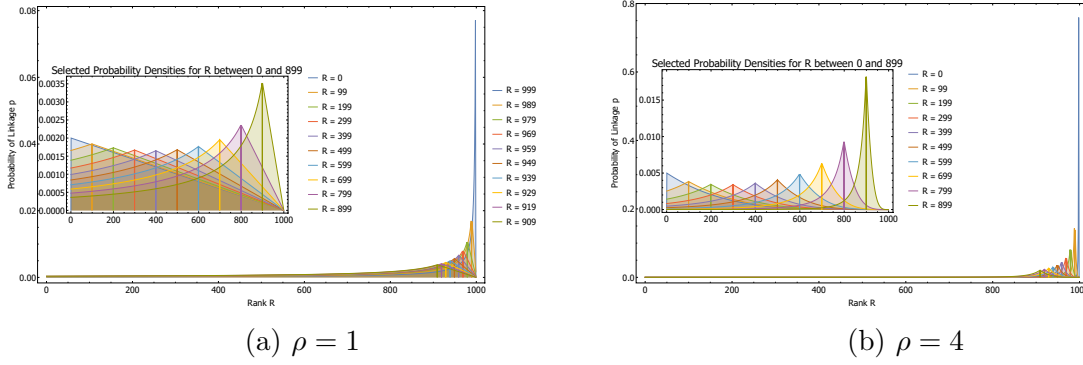


Figure 6.2: Theoretical Probability Density Functions (PDFs) of a node with a given income rank  $R$  for linkage with another node for the whole support of income ranks.

Overall, an agent’s choice of link neighbours depends on the homophily level  $g$  and the wage gap  $\rho$ . The resulting network is a member of the family of Random Geometric Graphs (Dall and Christensen 2002), which Talaga and Nowak (2020); Schulz et al. (2022) showed to reproduce core features of social networks efficiently, especially regarding their small-world character (Watts and Strogatz 1998). These include, in particular, a rather short average path lengths and high degrees of clustering. Specifically, the model combines the concept of homophily (Boguná et al. 2004) with pre-setting node degrees (Newman et al. 2001; Newman 2009). However, concerning our application, we are able to simplify both approaches by pre-determination of only the global minimum degree, like in Preferential-Attachment networks, and consequently defining relative weights rather than absolute probabilities.

### 6.3.3 Individual Perception

Like the linkage, any evaluation is done *by* the agents themselves given their individual knowledge. This knowledge consists of a local component defined by an agent’s ego network and a global component, which is simply the true wage gap.

#### Comparison within One’s Neighbourhood

Agents know about their own income and also their social contacts’ incomes. However, within the baseline specification, they do not possess knowledge about any other agent or structural features of the whole income distribution. Let  $U$  and  $P$  be the set of

underprivileged and privileged agents and  $\Theta_i$  the perception set of agent  $i$  consisting of  $N_{U,i} \in \mathbb{N}_0$  underprivileged and  $N_{P,i} \in \mathbb{N}_0$  privileged agents (including agent  $i$  themself). Agent  $i$  then calculates the perceived local wage gap  $l_i$  as

$$l_i = \frac{\bar{Y}_P^i - \bar{Y}_U^i}{\frac{1}{2}(\bar{Y}_P^i + \bar{Y}_U^i)},$$

with  $\bar{Y}_P^i = (1/N_{P,i}) \cdot \sum_{j \in \{\Theta_i \cap P\}} Y_j, \bar{Y}_U^i = (1/N_{U,i}) \cdot \sum_{j \in \{\Theta_i \cap U\}} Y_j$

and  $N_{U,i} \neq 0$  as well as  $N_{P,i} \neq 0$ .

For either  $N_{U,i} = 0$  or  $N_{P,i} = 0$ , we set  $l_i = 0$ , as in this case, agent  $i$  is unable to observe any wage gap.  $l_i$  is thus the percentage difference in perceived mean wages of the underprivileged and privileged agents within the perception set of  $i$ . The perceived local wage gap corresponds to the wage gap perception within any agent's ego network. We opt for the particular approximation above because it is symmetric (Törnqvist et al. 1985) and bounded in  $[-2, 2]$  and thus more robust against outliers (Decker et al. 2014). Our results are not materially sensitive to the specific symmetric approximation of the growth rate and are very similar to the more common approximation by log-differences.<sup>6</sup>

### Processing of a Composite Signal

Informed by our initial simulation results and the empirical finding that individuals are often aware of the wage gap through, e.g., national media, we consider the possibility that agents process a composite signal that combines their local perception and the global average wage gap. In this parsimonious extension, the composite perception  $p_i$  of individual  $i$  is a linear combination of local perceptions  $l_i$  and the (correct) global wage gap  $g$  by

$$p_i = (1 - w_i) \cdot l_i + w_i \cdot g. \quad (6.3)$$

The only free parameter is thus  $w_i \in \mathbb{R}$ , the weight  $i$  puts on the global signal. It is straightforward to solve for  $w_i$  as

$$w_i = \frac{p_i - l_i}{g - l_i}, \quad (6.4)$$

with intuitive comparative statics. Whenever the composite perception  $p_i$  increases relative to  $l_i$ , the weight on the global signal increases, as local knowledge is less relevant for perceptions in these cases. Our dataset includes  $g$  and the mean perceptions for the underprivileged and privileged groups, i.e.,  $\bar{p}_U$  and  $\bar{p}_P$ . Of course, the aggregation function (6.4) in the form of a weighted average is not the only possible way to aggregate local and global signals. We chose this particular function form for its tractability and intuitive

---

<sup>6</sup>Results are available upon request.

interpretability. Here, the respective weights  $w$  and  $(1 - w)$  are immediately interpretable as percentage weights in contrast to more technically involved formulations.

Together with the simulation means for the local signal,  $\bar{l}_U$  and  $\bar{l}_P$ , we are able to estimate the implied mean population weights  $\bar{w}_U$  and  $\bar{w}_P$  by replacing the individual parameters with the population means. In Appendix 6.A, we show that using mean perceptions delivers exact results for the implied mean population weights, whenever the weights  $w$  within a subpopulation are uncorrelated to the local perceptions of the this group or, whenever the weights within each group of privileged and underprivileged individuals are homogenous, i.e.,  $\bar{w}_U$  is the same constant weight for all underprivileged agents and  $\bar{w}_P$  is equal for all privileged agents. We consider deviations from this identification assumption in the discussion section below.

### Perceiving Overall Inequality

We also consider the perception of *overall* inequality, i.e., the Gini coefficient rather than *intergroup* inequality as the wage gap between privilege groups. The local Gini perception mirrors the standard calculation of the Gini coefficient but is applied to the set of individuals  $i$  observes  $\Theta_i$  and themselves. The local Gini  $L_i$  is thus defined as

$$L_i = \frac{\sum_{j \in \Theta_i} \sum_{k \in \Theta_i} |Y_j - Y_k|}{2n^{-1} \sum_{j \in \Theta_i} Y_j}. \quad (6.5)$$

This local or perceived Gini coefficient  $L$  equals the ratio of the sum of all differences of incomes within the ego network of  $i$  (including agent  $i$ 's income) to twice the mean income within this ego network (again calculated including agent  $i$ 's income). Since the functional form of this calculation is equivalent to the standard Gini formula, bias only arises whenever the ego network is unrepresentative of the total population. This is the case for  $\rho > 0$ , with ‘unrepresentativeness’ increasing in  $\rho$ . For  $\rho = 0$ , perceptions are unbiased on average and irrespective of privilege class, as we also verify by simulation.

### 6.3.4 Validation Strategy

To ensure internal validity of our model, i.e., detect any potential bugs, we obtained not only data on the wage gap but also general inequality perception and the network topology. This data is in line with what Schulz et al. (2022) find for their model without a wage gap, meaning that our model features the mechanisms that we intend. These mechanisms are to be understood as a (possible) minimal set of necessary assumptions (Grüne-Yanoff 2009) to explain the empirically observed underestimation of wage gaps. Hence, our model presents a how-possibly explanation looking for mechanisms that could potentially cause the observed phenomenon and does not claim that these conditions are fulfilled in the real world. Nevertheless, there is a resemblance (Mäki 2009) between our model and the real



world as the “model produces quantitative agreement with empirical macrostructures, as established through on-board statistical estimation routines” and also “quantitative agreement with empirical microstructures” (Barde and Van Der Hoog 2017). We use empirical micro-data to calibrate our model, namely an exponential income distribution and wage gaps that characterise industrialised countries as well as a linking function that is plausibly micro-founded within the discrete-choice approach. Furthermore, we compare the simulation output to available empirical macro-data. Therefore, the explanans can be true in the real world and the cause for the observed empirical fact: Our proposed mechanism fulfils the minimum conditions for a good epistemically possible how-possibly explanation formulated by Grüne-Yanoff and Verreault-Julien (2021). That our simulation results are in “qualitative agreement with empirical macrostructures” (Fagiolo et al. 2019, p. 771), namely successful replication of the stylised empirical facts, affirms the external validity of the model. Thus, the model is both technically verified as well as validated based on input and output measures. This combination of internal and external validity is a rare but desirable feature of simulation models (Gräbner 2018).

## 6.4 Results

We use  $N = 1,000$  agents for each simulation run that are equally split into a ‘privileged’ and ‘underprivileged’ class which thus contain 500 members each. 1,000 agents ensure computability but is also in line with the sample sizes of empirical studies into wage distributions (Tao et al. 2019).<sup>7</sup> Both the income distribution for the privileged and underprivileged class are initialised by an identical exponential distribution that is characterised by a rate parameter  $\lambda = 1$ , i.e., mean income is also exactly equal to unity. All incomes within the distribution of the underprivileged class are then downscaled by the wage gap  $(1 - g)$ , as indicated by the empirical results in Shaikh et al. (2014). For our calibration by sample moments in subsection 6.4.3, we use the empirical estimates by Malul (2021), again assuming a total population  $N = 1,000$ . Since our network generation mechanism is partly stochastic, our presented results are generally averages over 100 Monte Carlo (MC) runs each if not otherwise indicated.

### 6.4.1 Mean Local Perceptions

We first look at the mean local perceptions according to privilege and varying  $g$  and  $\rho$  in the empirically relevant ranges (without impact of a global signal). Recall that the scarce empirical literature suggests that (i) the wage gap is underestimated regardless of

---

<sup>7</sup>Moreover, Schulz et al. (2022) carried out sensitivity analyses with higher  $N$  and found all results to be robust. They also showed analytically that all main model mechanisms are independent of sample size.

privilege but that (ii) the underprivileged have much more accurate perceptions than the privileged. Figure 6.3 summarises this first battery of simulation results.

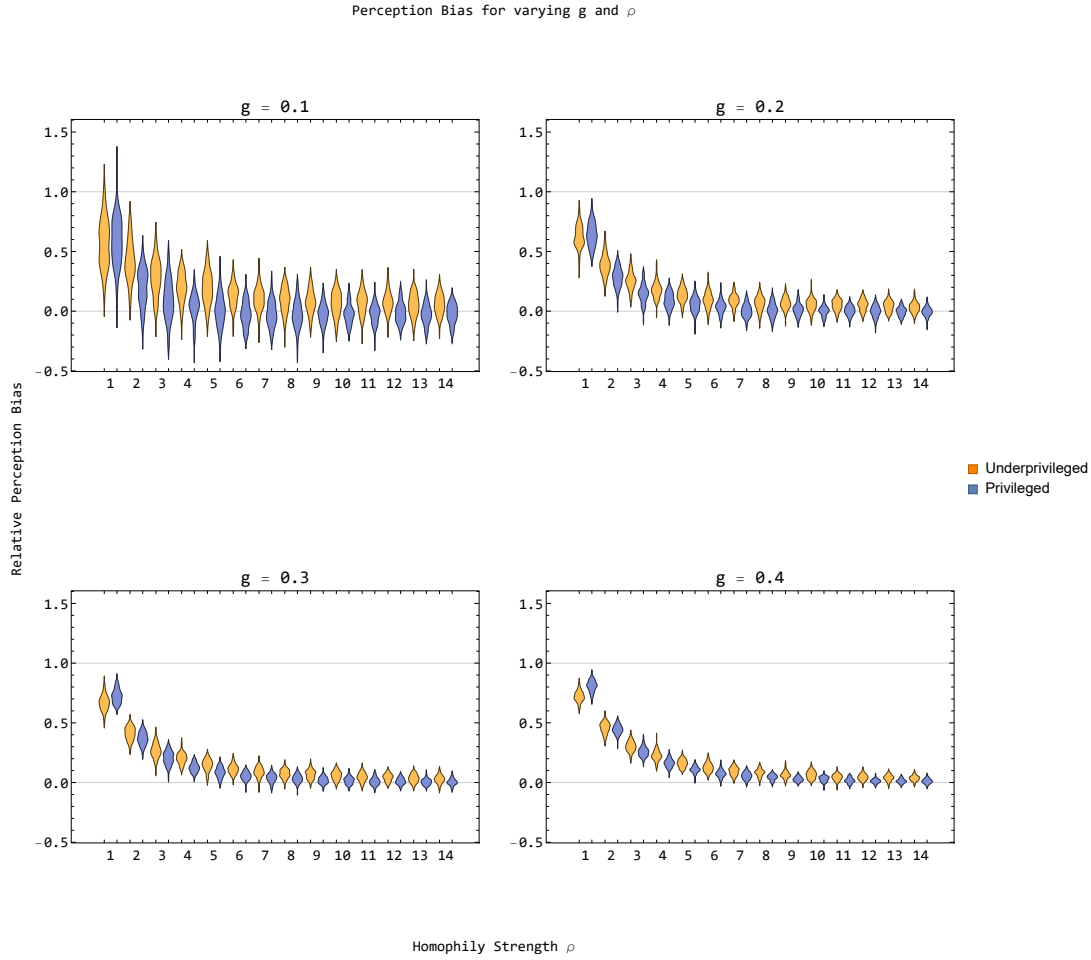


Figure 6.3: Violin plots of perceived wage gap  $p_i$  relative to the true wage gap  $g$  for various  $g$  and  $\rho$ .

In line with stylised fact (i), our simulated population overall tends to underestimate the wage gap considerably, regardless of privilege class, when we only consider the local perceptions. We find that (almost) all violin plots are significantly below the gridline at unity that would indicate correct local perceptions. Of course, this outcome is entirely unsurprising. All agents estimate the wage gap based on the set of neighbours that have similar wage levels, since the graph formation is homophilic in income. It is thus improbable for any agent to observe incomes that strongly deviate from their own income in question, including most notably incomes from a different privilege class. Even though there is a unidirectional downwards bias, this bias affects the two privilege classes differently. Since selectivity is locally maximal at the lower and upper tail of the income distribution, as we show in Figure 6.2, overall selectivity exhibits a U-shaped pattern. A wage gap  $g > 0$  in combination with income homophily by  $\rho > 0$  thus causes two counteracting effects within our model. First, since  $g > 0$  moves the incomes of the richest underprivileged agents

closer to the median income, it tends to decrease their selectivity, thus generally improving the accuracy of their wage gap estimation. Second,  $g > 0$  also pushed the incomes of the poorest underprivileged agents away from the incomes of the poorest privileged agents and lets the poor underprivileged inhabit the lower tail of the income distribution by themselves. Thus, the neighbourhoods of the underprivileged poor are on average less diverse and their estimates of the wage gap are downwards biased. The direction of the composite effect depends on both  $g$  and  $\rho$ , as Figure 6.3 shows. Generally, increasing  $\rho$  increases the relative strength of the first partial effect since the richest become much more selective here. Because our proposed mechanism is based entirely on income differences, increasing both  $g$  and  $\rho$  increases not only income segregation but also segregation based on privilege class. Conversely, the lower  $g$  and  $\rho$ , the more random the network generation will be and the higher the variability of simulation outcomes. As a stark illustration of this variability, note that the lowest  $g = 0.1$  and  $\rho = 1$  we consider is consistent with both an *overestimation* of the wage gap (simulation means above the gridline at unity) and with *reversed* estimates where agents believe the privileged to earn less on average, i.e., mean estimates reaching below the gridline at zero. This effect is evident from the fact that variability in mean perceptions decreases both in  $\rho$  within each panel and in  $g$  across panels in Figure 6.3.

We find that stylised fact (ii) is only possible in our model setup for homophily levels  $\rho \geq 2$ . This is in line with our findings for the perception of overall inequality, measured as the perceived Gini coefficient, that would imply  $\rho \in [4, 14]$ .<sup>8</sup> Note that homophily strengths in this range imply a tremendous downward bias in perceived wage gaps of at least 50%.

While we find that the two stylised facts (i) and (ii) can be *qualitatively* replicated for homophily strengths  $\rho \geq 2$ , our *quantitative* estimation results indicate that the downwards bias is much too high. In the sample by Malul (2021) with a true wage gap  $g$  close to  $g = 0.3$ , men underestimate the gender wage gap by about 36% and women only by about 22%. This is, of course, far from the at least about 50% downwards bias that is, given our simulation results, necessary for underprivileged agents to be on average closer to reality than the privileged, in line with stylised fact (ii) and also consistent with our back-of-the-envelope calibration for overall inequality perceptions. This might indicate that, in contrast to perceptions of *overall* inequality, perceptions of *gender* inequality might be formed as a combination of global and local signals. We calibrate our model in

---

<sup>8</sup>For this calibration, we calculate the mean perceived Gini coefficient for the whole sample population and vary  $\rho$  to replicate the maximum and minimum perceived Gini that Choi (2019) finds for his large sample from the International Social Survey Program. The sample covers 32 OECD countries and 30 sample years. Over all countries, he finds a minimum perceived Gini of  $G_{min} = 0.1276$ , a mean perceived Gini of  $G_{mean} = 0.1708$  and a maximum perceived Gini of  $G_{max} = 0.2534$ . The maximum Gini  $G_{max}$  translates to  $\rho \approx 4$ , while the minimum Gini  $G_{min}$  implies  $\rho \approx 14$ .

line with the composite rule in eq. (6.4) in subsection 6.4.3 to spell out the implications of this in more detail.

### 6.4.2 Disaggregating Local Perceptions

The relatively straightforward impact of income and homophily on the perceived wage gaps at the aggregate level masks substantial differences for individual agents. Here, *neighbourhood disassortativity*  $D$  presents the critical link between an agent's income and wage gap perception. It is the rate of link neighbours of the opposite privilege class. Hence,  $D_i = 0$  means that agent  $i$  is located within a segregated community of one's own privilege class whereas  $i$  being isolated among members of the other class results in  $D_i = 1$ ; a disassortativity slightly below 0.5 (the exact rate depending on the size of the ego network with 0.33 for 5 link neighbours) indicates equal numbers within the group. We expect disassortativity to be positively correlated to income for underprivileged agents: The wage gap decreases diversity among the poorest of them because they become isolated from the poorest privileged agents; concurrently, those underprivileged agents with a pre wage gap income higher than the global median now get pulled closer to it and experience a high diversity level (with disassortativity close to 0.5) due to the low selectivity there. Contrary to that, we expect a negative relationship between income and disassortativity when it comes privileged agents: The highest earners of them are also the richest overall and they experience few underprivileged agents in their link-neighbourhoods amounting to rather low disassortativity levels. By contrast, the wage gap pushes the poorest privileged closer to the the overall median income meaning more frequent encounters with underprivileged agents as explained above. One would expect these mechanisms to grow more powerful in  $g$ , for this results in a greater divergence of in-group income ranks and global ones. The homophily strength  $\rho > 0$  should work as another catalyst of the described mechanisms.

Figures 6.4 and 6.5 depict this relationship for examples of  $g$  and  $\rho$ .<sup>9</sup> A vertical section through a heatmap represents all levels of neighbourhood disassortativity - and resulting perceptions - that any specific individual agent (identified across MC runs by their income) for the respective combination of wage gap and homophily. Chance has a substantial impact on low income ranks, and the disassortativity varies largely between MC runs. Due to the linkage selectivity, the range of realised disassortativity levels shrinks for higher income ranks. That this effect becomes stronger with rising  $\rho$  testifies to the linkage procedure discussed in the section on the model setup. Thus, underprivileged agents with higher incomes only display high disassortativity levels, in line with the theoretical expectations above. However, while the range of realised disassortativity levels shrinks

---

<sup>9</sup>The figures present a selection of values for  $g$  and  $\rho$  that are empirically plausible. Yet, the data for other parameter combinations display the complex and non-monotonous patterns as well.

in income for the privileged, where this range lies for the highest incomes depends on  $g$ . Furthermore, the relationship between income and disassortativity is highly non-linear.

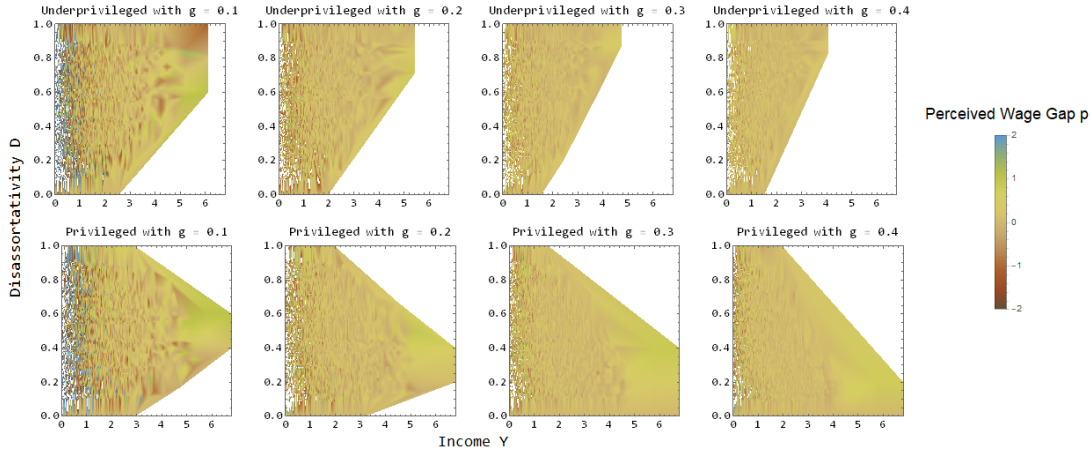


Figure 6.4: Heatmap of wage gap perceptions (as a ratio to the correct wage gap  $g$ ) for varying privilege, disassortativity  $D$  (rate of links to the other privilege class) and income  $Y$  for  $\rho = 8$  and varying actual wage gaps  $g \in \{0.1; 0.2; 0.3; 0.4\}$ . Upper panels for the underprivileged agents, lower panels for the privileged ones.

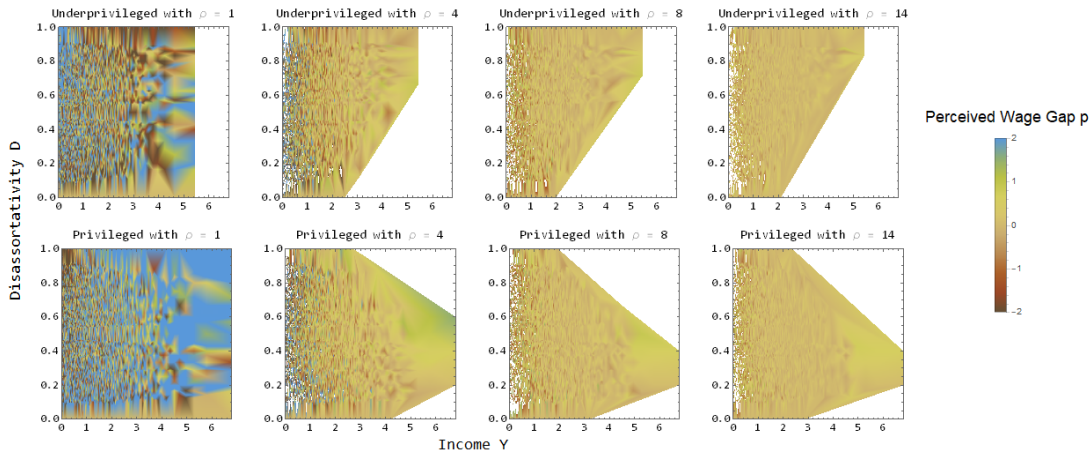


Figure 6.5: Heatmap of wage gap perceptions (as a ratio to the correct wage gap  $g$ ) for varying privilege, disassortativity  $D$  (rate of links to the other privilege class) and income  $Y$  for  $g = 0.2$  and varying actual wage gaps  $\rho \in \{1; 4; 8; 14\}$ . Upper panels for the underprivileged agents, lower panels for the privileged ones.

Non-linear and even non-monotonous patterns also surface in the effects that a rising  $\rho$  for a given  $g$  - or vice versa - have on disassortativity. An agent's overall income rank changes in  $g$ , and the underprivileged agents with the highest income can grow less disassortative for very high  $g$  as Figure 6.4 illustrates. Figure 6.5 highlights that even moderate homophily levels significantly impact disassortativity, especially of privileged agents. Nevertheless, while an increase in  $\rho > 0$  from low to moderate levels decreases the range of realised disassortativity levels, this range may expand again for some agents when  $\rho$  gets high.

An agent's different neighbourhood disassortativity levels themselves affect this agent's local wage gap perception ( $l$ ). With this, Figures 6.4 and 6.5 testify to the interaction of  $Y$  and  $D$  (both values being themselves traced to the pre wage gap distribution,  $g$ , and  $\rho$ ) and its complex impact on  $l$ . Namely, they point to three obvious phenomena:

- For a given  $\rho$ , a higher actual wage gap  $g$  does not imply a higher perceived wage gap for all agents: Trivially, an agent's perception depends on their link-neighbours but a higher actual wage gap for the whole population can mean shrinking local differences between privileged and underprivileged members of some ego networks.
- Given any income and income rank distribution implied by the wage gap, rising homophily lowers the average wage gap perceptions overall as expected. However, some agents'  $l$  behaves non-monotonically and may rise in  $\rho$ .
- Chance has a significant impact on many agents' ego networks (i.e., their disassortativity) and the resulting perceptions, as the great variety in  $D$  and  $l$  between MC runs - for vertical sections through heatmaps - demonstrates.

As a corollary of these phenomena, wage gap perception differs from general inequality one in Schulz et al. (2022), one cannot predict an agent's individual wage gap estimate  $l$  by their income rank; and their privilege class does not work as a predictor, too. It is neither  $Y$  nor  $D$  alone that determines an agents' wage gap perception but the interaction of the two. Given sufficiently high levels of  $g$  and  $\rho$ , paints a clear and monotonous picture for high income ranks and at the edges of the realised disassortativity levels. However, for the majority of agents and MC runs (i.e. disassortativity levels of an agent with a given income rank), there is no such monotonous behaviour of  $l$ .

These behaviours occur because the likeliness of an agent with a given income rank in their privilege class to be picked as another agent's link-neighbour simultaneously depends on  $g$  and  $\rho$ , as shown in Section 6.3.2. The wage gap determines the likeliness ranking of agents as link neighbours. If for one agent the relative likeliness ranks of only two other agents, who belong to different privilege classes, change, all three agents may experience notably changed wage gap perceptions. The homophily level specifies how much the likelihood of selection as link neighbours differs between agents. Hence, the perception landscapes become more 'rugged' when  $\rho$  increases. Moreover, like disassortativity, it depends on income how much the perceptions vary between MC runs. Thereby, agents with the highest incomes exhibit the least variability because they are the most selective regarding their link-neighbours, giving less room for chance.

### 6.4.3 Empirical Calibration and Composite Signal

The simulation results in subsection 6.4.1 indicate that the effect strength from homophilic segregation on local perceptions is too strong to replicate the empirically observed effect

sizes in Malul (2021). To account for this, we extend our rule for information processing along the lines introduced in subsection 6.3.3 to allow for a *composite signal* aggregated from global and local information, respectively. We calibrate our baseline model with the empirically observed parameters to deduce the weight both the privileged (male) and the underprivileged (female) class puts on the global signal. To achieve this, we use a ‘naive estimator’, where we replace the individual  $p_i$  and  $l_i$  in eq. (6.4) with the respective sample means  $\bar{p}$  and  $\bar{l}$ . This estimator is exact, whenever the individual weights  $w_i$  that individuals  $i$  place on the global signal are uncorrelated to local perceptions  $l_i$ , as we show in Appendix 6.A. The empirical gender wage gap for the Israeli sample is  $g = 0.317$ , with a female labour force participation rate of 0.472 and a male one of approximately 0.528. The target mean perception of the underprivileged is  $\bar{p}_U = 0.292$  and of the privileged  $\bar{p}_P = 0.203$ . We initialise our model with a total population of  $N = 1,000$  which implies an underprivileged population size of  $N_U = 472$  and a privileged population size of  $N_P = 528$ . The estimated implied mean weights  $\bar{w}_U$  and  $\bar{w}_P$  for both the underprivileged and privileged class and different  $\rho$  are shown in Figure 6.6.

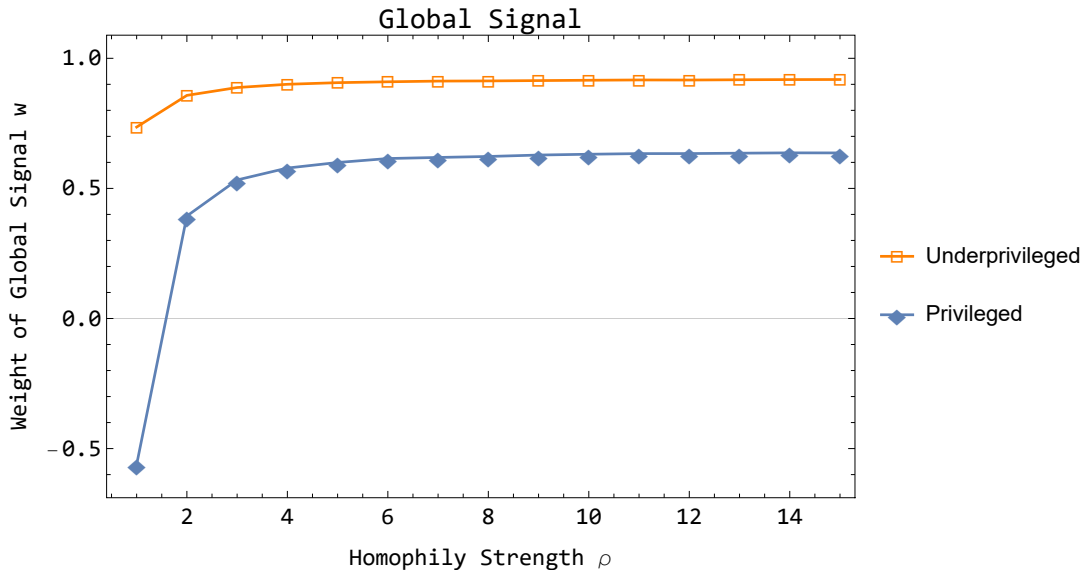


Figure 6.6: Implied weights for the global signal for both privilege classes and the empirical calibration.

For the whole range of estimates, we find that the underprivileged group places a much higher weight on the global signal than the privileged one. In the (empirically implausible) case of  $\rho = 1$ , the implied weight of the privileged even becomes negative, which would amount to the privileged part of the population actively discarding global information in favour of ‘overweighting’ their biased local signals. The gap between the underprivileged and privileged classes becomes more narrow with higher homophily when the local perceptions of the underprivileged decrease but continues to be sizeable for all considered  $\rho$ . Note that this also implies that the composite perceptions of the underprivileged are much less volatile, as they put more weight on the (correct) global signal rather than the

(noisy) local one, which is evident from the exemplary violin plots for  $\rho = 8$  in Figure 6.7. In this case of  $\rho = 8$ , the difference in the variance of estimates is more than one order of magnitude, with the variance of mean composite perceptions of the privileged being more than tenfold the variance of mean composite perceptions of the underprivileged.

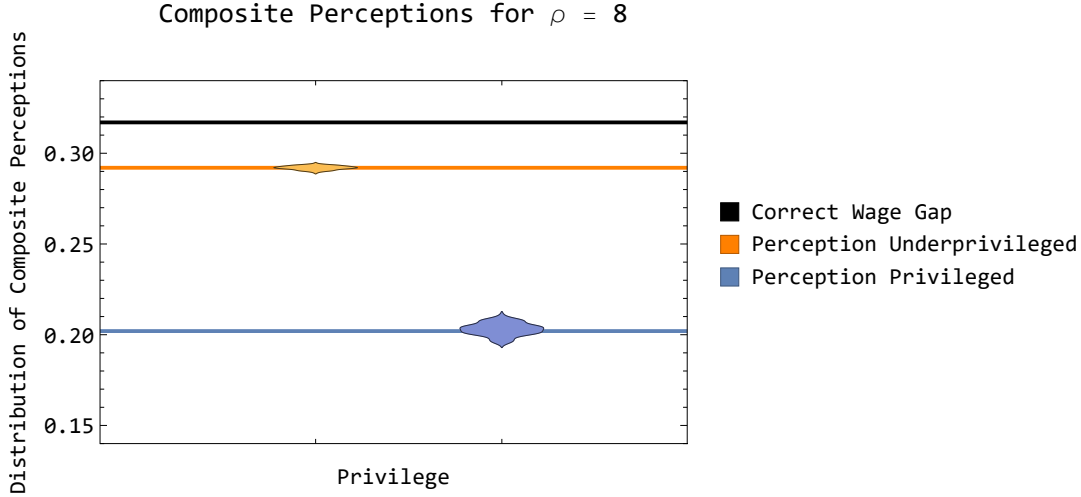


Figure 6.7: Violin plots for the composite perceptions of privileged and underprivileged agents for the empirical calibration and  $\rho = 8$ .

Our calibration exercise thus yields two main results: Firstly, we find that local perception formation on skewed information sets is insufficient to generate the observed empirical effect sizes. Only for composite signals that combine both a (correct) global signal with the local perception can the model be reconciled with the empirical evidence provided in Malul (2021). By contrast, the local signal is sufficient for perceptions of *global inequality* (Schulz et al. 2022), hinting at conceptual differences between both. They are important because they might imply that the concept of a ‘wage gap’ can be much more easily conveyed within educational campaigns than the arguably much more complex and ambiguous concept of ‘inequality’. Indeed, the empirical evidence suggests that concept complexity and the associated cognitive costs might exhibit significant effects on individual behaviour (Oprea 2020). Secondly, we find that the underprivileged place much higher weight on the global signal than the privileged, with much less noisy estimates. This is consistent with the empirical finding that the adversely affected part of the population is more interested in global information about the issue, as is also documented on the micro level within the empirical psychological literature (Wu 2021).

Our results from calibrating the composite perception rule suggest that belief formation regarding intergroup inequality is based not only on local but also on global information. Since our proposed mechanism is purely based on economic considerations, this also indicates that the implied differential weights men and women place on the global signal requires an explanation by factors outside of the network-segregation mechanism we propose.



### 6.4.4 The Impact of Privilege on General Inequality Perceptions

In contrast to perceptions of the wage gap, the ones of biases regarding overall rather than gender-based inequality can be generated from network segregation only, as Figure 6.8 demonstrates. For all considered  $\rho > 0$  and  $g > 0$ , the underprivileged appear to observe a greater degree of economic inequality within the population. The perceived degree of inequality obviously decays in  $\rho$  as network segregation increases. Hence, the introduction of privilege classes to the model does not alter its outcomes concerning general inequality perceptions described by Schulz et al. (2022). The relative bias of the privileged increases in  $g$ , that is, the underprivileged become relatively more accurate in their perceptions compared to the privileged, the higher the wage gap is. Paradoxically, being economically underprivileged thus translates into privileged epistemic access regarding the actual state of inequality in our model.

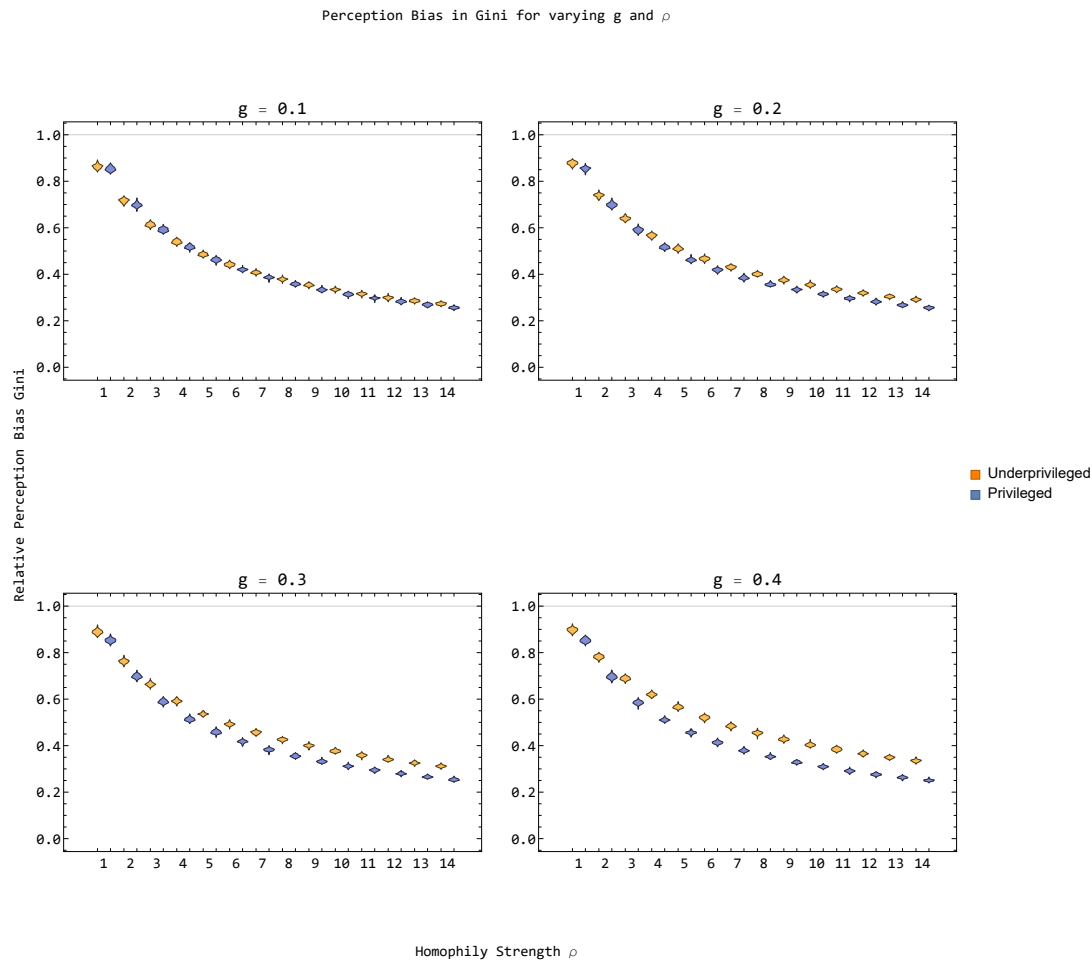


Figure 6.8: Violin plots of (normalised) perceived Gini coefficients for various  $\rho$  and  $g$  according to privilege class. Perceptions are displayed in relation to the actual Gini of  $G \approx 0.5$ .

In contrast to the vast literature on redistributive preferences according to race and gender surveyed in section 6.2, this finding constitutes an emergent outcome of the homophily in

ego network formation. As Schulz et al. (2022) note, the homophilic graph formation lets the income differences in the numerator of eq. (6.5) increase linearly at most in income rank. In contrast, the mean perceived income in the denominator increases much faster exponentially due to the underlying exponential wage distribution. As a result, individuals with lower income (ranks) perceive higher Gini coefficients  $L$  and are more accurate on average. Pushing the underprivileged to the lower income ranks by the wage gap  $g$  thus increases their average perception. Figure 6.9 highlights this effect, with perceived Ginis  $L$  visibly decaying in income rank. As is readily apparent from the same figure, the density of underprivileged agents (with orange markers) is much higher at lower ranks of the rank distribution, indicating that those also perceive higher levels of inequality on average.

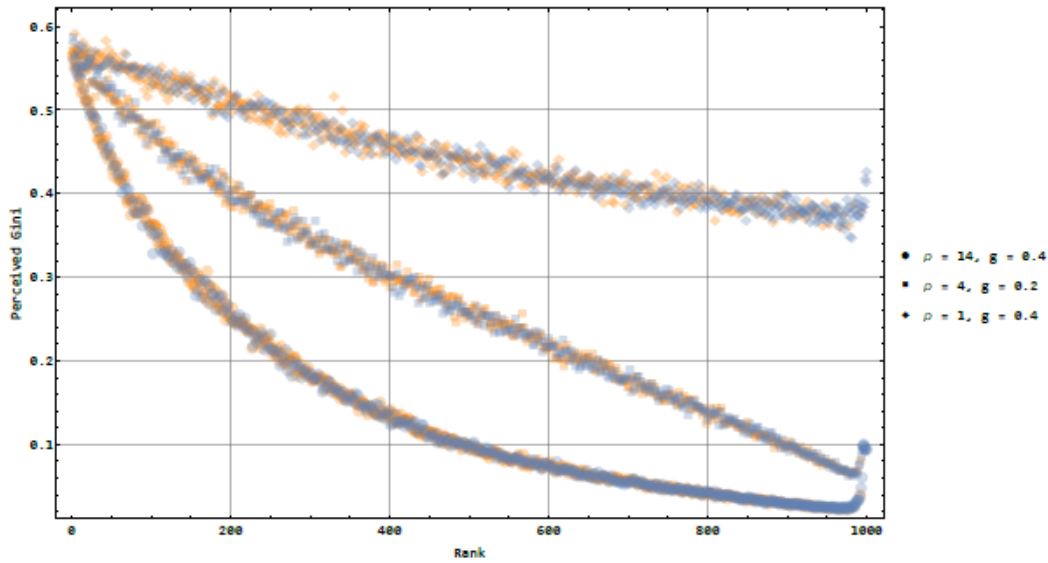


Figure 6.9: Perception of the local Gini coefficients per income rank (increasing order) and according to privilege class (orange for the underprivileged, blue for the privileged).

We are thus able to replicate the finding that perceptions anchored within ego networks ( $L$ ) are sufficient to generate the empirically observed patterns of overall inequality perceptions. Our findings also directly suggest that it is not necessarily non-economic factors like normative beliefs or self-interest driving the gender and racial differentials in support of redistributive policies. Instead, it might simply be the case that skewed samples from homophilic ego networks render the privileged particularly ignorant regarding issues of economic inequality, as our simulation results would suggest.

## 6.5 Discussion

The agent-based simulation model presented in this paper suggests that the mild underestimation of gender and racial wage gaps, which empirical studies describe, results from individuals processing information obtained by combining one's link-neighbours' incomes

with a global signal reflecting the actual wage gap. However, underprivileged agents put considerably higher weights on the global signal than privileged ones. Therefore, those adversely affected by a wage gap may perceive it more severely, i.e., closer to its actual severity, simply because they listen more carefully to global information. Likewise, underprivileged agents underestimate income inequality, in general, less strongly than privileged ones because of their lower selectivity when choosing link neighbours, the latter itself being a corollary of an agent's own income level and rank. Different from wage gap perception, a global signal is not necessary to explain the empirical underestimation of general wage inequality. Furthermore, the level of underestimation hereby depends solely on an individual's income (with lower incomes typically meaning a less severe underestimation) - not their privilege class.

We achieve these conclusions by imposing a rather strong identification assumption, i.e., local perceptions and weights for the two signals being uncorrelated or homogeneous within groups. This leads to a 'naive estimator' for the implied weights for local and global signals, where we simply substitute the mean local perceptions  $\bar{l}_U$  and  $\bar{l}_P$  for the underprivileged and privileged in eq. (6.4) that is stated in terms of the individual local perceptions  $l_i$ . Fortunately, we recover the mean weights of the respective subpopulations  $\bar{w}_U$  and  $w_P$ , whenever the individual cross-sectional weights and local perceptions are uncorrelated for the given subpopulation or whenever weights are homogeneous in them, as we show in Appendix 6.A. For a positive correlation between individual weights of the global signal and local perceptions, we underestimate the true mean and for a negative correlation, we overestimate the mean weight with our naive estimator. One could make a plausible case for both types of correlation: Either high local perceptions let individuals pay more attention to their immediate lived experience, which would imply a negative correlation or local perceptions of high inequality prime people into being more responsive to global signals about the issue which would then imply a positive relationship between both. Since we do not have any information on the correlation structure within the sample by Malul (2021), this assumption of zero correlation appears to be the most natural one to us, then directly implying the naive estimator. For further applied studies, working out the actual correlational patterns should, however, be a priority and might shed light on further theoretical implications of the two correlation mechanisms we sketched, i.e., correlation patterns differing between privileged and underprivileged groups, therefore, explaining part of the difference in estimated implied weights.

For policy-makers and the public debate, this finding suggests that underprivileged groups are a reliable source, and their account of the wage gap is not strongly biased by subjective experience or even self-interest but on a higher emphasis on what is objectively the case. Furthering such emphasis on the globally true value, e.g. through education treatments, seems to be especially promising in the context of a wage gap that is easy to comprehend. These measures should especially target the privileged who put lower weight on the global

signal and rely more on the biased conclusions that they draw from observing their ego networks. Another challenge raised by our model is that income homophily can perpetuate homophily in privilege class (e.g., racist self-segregation) even though individuals do not willingly segregate by these privilege groups. Thus, when taking any actions to combat such segregation, policy makers should observe income homophily, too.<sup>10</sup>

In many ways, the global signal and the associated weights remain a residual quantity within our model that is not further explained. In obvious analogy to the concept of “total factor productivity” as a growth residual in economics, the global weights we identify are thus more than anything “measure[s] of our ignorance” Abramovitz (1993) concerning the actual underlying cognitive processes. However, we hope that our taxonomy of global and local signals as well as the mechanism for composite perception formation informs further empirical research on determinants of the global weight. In particular, we identified the correlational structure between local perceptions and global weights to be of potentially immense importance in determining composite perceptions.

We find that individual level phenomena react highly sensitively to both predetermined and random initial conditions. This property potentially poses obstacles to applied empirical research - given resemblance between our model network and the real world one. The complexity of the landscape indicates that average population findings mask important heterogeneity on an individual level; thus there is no simple meaningful aggregate representation. Moreover, our findings indicate that the relationships between income/actual wage gap, disassortativity, and perceptions do not only depend on the actual wage gap, the strength of homophily and the privilege class. Instead, the ego network disassortativity and perception vary along the income distribution - and the steepness of this variability is also non-monotonous. Therefore, it is impossible to fit any parsimonious monotonic function to any of these relationships.<sup>11</sup> Per construction, the relationships are deterministic in our model, but in the presence of the sketched non-monotonicities, empirical work might fail to detect the underlying structures, relationships and their causal mechanisms we impose. The striking non-monotonicities in individual perceptions imply a need for great caution when employing standard statistical tools in such further research, though. Namely, Pearson correlations are not meaningful since the observed relationships are highly case-dependent. Moreover, standard measures of non-linear relationships like Kendall’s  $\tau$  or Spearman’s  $\rho$  do not work in the studied situation of two groups that react differently to a change in an external variable. Finally, individual perceptions depend on the specific network layout. In turn, this layout emerges from the interplay of the actual wage gap, homophily, and chance. Hence, reliable predictions are

---

<sup>10</sup>We thank an anonymous reviewer for raising this last policy implication.

<sup>11</sup>Indeed, the Pearson correlation coefficients between local perceptions  $l$  and disassortativity  $D$  for all simulation runs and considered parameter combinations do not exceed 0.1 in absolute value.

impossible at an aggregate level and instead require investigation of individuals in their network interactions.

Our general framework points to obvious theoretical extensions: First, we assumed a common  $\rho$  for all individuals throughout all experiments, irrespective of their privilege class as the most parsimonious assumption. A natural extension would allow for heterogeneity in  $\rho$  across groups to examine differential homophily. Also, we only considered two groups each so far but neglected that identity is an inherently multidimensional construct with discrimination and marginalisation operating on different layers. Therefore, model extensions should consider the interaction of both racial and gender-based discrimination on individual perceptions. An *intersectional* perspective should account for the fact that wage gaps are *superadditive* (Stelzner and Bahn 2022; Bright et al. 2016). Apart from considering more than two groups, exploring differing functional forms for the superadditivity of wage gaps will surely also provide interesting and promising new avenues for further research.

## 6.A Appendix: Calibrating the Composite Wage Gap

Since we only observe the population means in our sample and especially are not able to determine any cross-sectional correlation between our variables, we impose some conditions on our estimator. In particular, we use the following expression as our naive estimator  $\hat{w}$  for the implied weight a subpopulation (the privileged or underprivileged) places on the global signal, where we replace the individual  $p_i$  and  $l_i$  in eq. (6.4) with the respective sample means  $\bar{p}$  and  $\bar{l}$ , i.e.,

$$\hat{w} = \frac{\bar{p} - \bar{l}}{g - \bar{l}}, \quad (6.6)$$

where  $\bar{p}$  corresponds to the mean perceived wage gap and  $\bar{l}$  the mean local signal of the given sampled subpopulation as well as  $g$  as the true global wage gap. Assume that all individuals form perceptions according to our composite perception rule. The estimator  $\hat{w}$  then corresponds to the true mean weight of the sampled subpopulation, whenever either i)  $w$  is equal for all considered individuals within the subpopulation or ii)  $w$  and  $l$  and thus  $(1 - w)$  and  $l$  are uncorrelated. To see this, consider the composite perception rule for a given individual  $i$ :

$$p_i = w_i \cdot g + (1 - w_i) \cdot l_i. \quad (6.7)$$

Averaging over all individuals  $i$  within the considered subpopulation of size  $n$  yields

$$n^{-1} \cdot \sum_{i=1}^n p_i = n^{-1} \cdot \sum_{i=1}^n w_i \cdot g + n^{-1} \cdot \sum_{i=1}^n (1 - w_i) l_i. \quad (6.8)$$

Expressing this in the bar notation from above in eq. (6.6), we get

$$\bar{p} = \bar{w}g + n^{-1} \cdot \sum_{i=1}^n (1 - w_i) l_i. \quad (6.9)$$

From expression (6.7), it becomes obvious that there is one degree of freedom left in our specification: The relationship between the vector of  $(1 - w)_{i=1}^n$  and  $(l)_{i=1}^n$ . We consider two cases:

- i) Assume first that  $w_i = \bar{w}, \forall i = 1, 2, \dots, n$ , that is, all individuals place equal weights on the global signal. This also implies that  $n^{-1} \cdot \sum_{i=1}^n (1 - w_i) = (1 - \bar{w})$  and thus allows to rewrite (6.7) as

$$\hat{w} = \frac{\bar{p} - \bar{l}}{g - \bar{l}}, \quad (6.10)$$

$$\bar{p} = \bar{w}g + (1 - \bar{w})\bar{l} \quad (6.11)$$

which is the desired result and gives rise to  $\bar{w} = \hat{w}$  according to eq. (6.6).

- ii) Assume now that the sample correlation between  $l$  and  $(1 - w)$ ,  $r_{1-w,l}$  exists and is exactly zero. Recall the definition of a sample correlation as

$$r_{wl} = \frac{\sum_{i=1}^n (1 - w_i) l_i - n \cdot \overline{(1 - w)} \cdot \bar{l}}{n s'_{1-w} s'_l} = \frac{\sum_{i=1}^n (1 - w_i) l_i - n \cdot (1 - \bar{w}) \cdot \bar{l}}{n s'_{1-w} s'_l}, \quad (6.12)$$

due to  $\overline{(1 - w)} = (1 - \bar{w})$  as well as  $s'_l > 0$  and  $s'_{1-w} > 0$  as the (corrected) sample standard deviation, i.e., both vectors exhibiting some variability. Imposing  $r_{1-w,l} = 0$  implies immediately that

$$\begin{aligned} \sum_{i=1}^n (1 - w_i) l_i &= n \cdot (1 - \bar{w}) \cdot \bar{l} \quad \text{and thus} \\ n^{-1} \sum_{i=1}^n (1 - w_i) l_i &= (1 - \bar{w}) \cdot \bar{l}. \end{aligned} \quad (6.13)$$

This again allows us to rewrite equation (6.7) as

$$\bar{p} = \bar{w}g + (1 - \bar{w})\bar{l} \quad (6.14)$$

which is the desired result and gives rise to  $\bar{w} = \hat{w}$  according to eq. (6.6).

The estimator is thus only exact, whenever either the correlation between  $(1 - w)_{i=1}^n$  and  $(l)_{i=1}^n$  is exactly zero or if all individuals place exactly the same weights on the global and local signals, respectively. Whenever  $r_{1-w,l} > 0$ ,  $\hat{w} > \bar{w}$  and we thus overestimate the true population mean (and vice versa). This follows straightforwardly from eq. (6.7). Solving for  $\bar{w}$ , the equation implies that, generally,

$$\bar{w} = \frac{\bar{p} - n^{-1} \cdot \sum_{i=1}^n (1 - w_i) l_i}{g}. \quad (6.15)$$

The estimator implies

$$\hat{w} = \frac{\bar{p} - (1 - \bar{w}) \bar{l}}{g}. \quad (6.16)$$

Whenever  $r_{1-w,l} > 0$  it follows that  $n^{-1} \cdot \sum_{i=1}^n (1 - w_i) > (1 - \bar{w}) \bar{l}$  by eq. (6.12) which implies by eqs. (6.15) and (6.16) that  $\bar{w} < \hat{w}$ . Conversely,  $r_{1-w,l} < 0$  implies by the same token that the inequality reverses and  $\bar{w} > \hat{w}$ , since  $n^{-1} \cdot \sum_{i=1}^n (1 - w_i) < (1 - \bar{w}) \bar{l}$ . Note finally that since  $r_{1-w,l} < 0$  implies that  $r_{w,l} > 0$  must hold,  $r_{w,l} > 0$  implies that  $\bar{w} > \hat{w}$  (and vice versa).

## References

- Abramovitz, M. (1993). The search for the sources of growth: areas of ignorance, old and new. *Journal of Economic History*, 53(2):217–243.
- Akchurin, M. and Lee, C.-S. (2013). Pathways to empowerment: Repertoires of women’s activism and gender earnings equality. *American Sociological Review*, 78(4):679–701.
- Akerlof, G. A. and Yellen, J. L. (1990). The fair wage-effort hypothesis and unemployment. *Quarterly Journal of Economics*, 105(2):255–283.
- Alesina, A., Ferroni, M. F., and Stantcheva, S. (2021). Perceptions of racial gaps, their causes, and ways to reduce them. *NBER Working Paper Series*, 29245.
- Alesina, A. and Giuliano, P. (2011). Preferences for redistribution. In Benhabib, J., Bisin, A., and Jackson, M. O., editors, *Handbook of Social Economics*, volume 1, pages 93–131. Elsevier, Amsterdam.
- Andreoni, J. and Vesterlund, L. (2001). Which is the fair sex? Gender differences in altruism. *Quarterly Journal of Economics*, 116(1):293–312.
- Auspurg, K., Hinz, T., and Sauer, C. (2017). Why should women get less? Evidence on the gender pay gap from multifactorial survey experiments. *American Sociological Review*, 82(1):179–210.
- Banerjee, A. and Yakovenko, V. M. (2010). Universal patterns of inequality. *New Journal of Physics*, 12(075032).

- Banerjee, A., Yakovenko, V. M., and Di Matteo, T. (2006). A study of the personal income distribution in Australia. *Physica A: Statistical Mechanics and its Applications*, 370(1):54–59.
- Barde, S. and Van Der Hoog, S. (2017). An empirical validation protocol for large-scale agent-based models. *Bielefeld Working Papers in Economics and Management*.
- Bavetta, S., Li Donni, P., and Marino, M. (2019). An empirical analysis of the determinants of perceived inequality. *Review of Income and Wealth*, 65(2):264–292.
- Boguná, M., Pastor-Satorras, R., Díaz-Guilera, A., and Arenas, A. (2004). Models of social networks based on social distance attachment. *Physical Review E*, 70(056122).
- Bright, L. K., Malinsky, D., and Thompson, M. (2016). Causally interpreting intersectionality theory. *Philosophy of Science*, 83(1):60–81.
- Carter, E. R., Brady, S. T., Murdock-Perriera, L. A., Gilbertson, M. K., Ablorh, T., and Murphy, M. C. (2019). The racial composition of students’ friendship networks predicts perceptions of injustice and involvement in collective action. *Journal of Theoretical Social Psychology*, 3(1):49–61.
- Chandra, A. (2000). Labor-market dropouts and the racial wage gap: 1940–1990. *American Economic Review*, 90(2):333–338.
- Choi, G. (2019). Revisiting the redistribution hypothesis with perceived inequality and redistributive preferences. *European Journal of Political Economy*, 58:220–244.
- Cusack, T., Iversen, T., and Rehm, P. (2006). Risks at work: The demand and supply sides of government redistribution. *Oxford review of economic policy*, 22(3):365–389.
- Dall, J. and Christensen, M. (2002). Random geometric graphs. *Physical Review E*, 66(016121).
- Davidai, S. and Walker, J. (2021). Americans misperceive racial disparities in economic mobility. *Personality and Social Psychology Bulletin*, 48(5):793 – 806.
- Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J. (2014). The role of entrepreneurship in US job creation and economic dynamism. *Journal of Economic Perspectives*, 28(3):3–24.
- Derzsy, N., Nédá, Z., and Santos, M. (2012). Income distribution patterns from a complete social security database. *Physica A: Statistical Mechanics and its Applications*, 391(22):5611–5619.
- Dos Santos, P. L. (2017). The principle of social scaling. *Complexity*, 8358909.
- Dragulescu, A. and Yakovenko, V. M. (2001a). Evidence for the exponential distribution of income in the USA. *European Physical Journal B*, 20(4):585–589.
- Dragulescu, A. and Yakovenko, V. M. (2001b). Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States. *Physica A: Statistical Mechanics and its Applications*, 299(1-2):213–221.



- Fagiolo, G., Guerini, M., Lamperti, F., Moneta, A., and Roventini, A. (2019). Validation of agent-based models in economics and finance. In Beisbart, C. and Saam, N. J., editors, *Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives*, pages 763–787. Springer International Publishing, Cham.
- Falk, A. and Knell, M. (2004). Choosing the Joneses: Endogenous goals and reference standards. *Scandinavian Journal of Economics*, 106(3):417–435.
- Fixman, M., Berenstein, A., Brea, J., Minnoni, M., Travizano, M., and Sarraute, C. (2016). A Bayesian approach to income inference in a communication network. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 579–582. IEEE.
- Foley, S., Kidder, D. L., and Powell, G. N. (2002). The perceived glass ceiling and justice perceptions: An investigation of Hispanic law associates. *Journal of Management*, 28(4):471–496.
- Franke, R. and Westerhoff, F. (2017). Taking stock: A rigorous modelling of animal spirits in macroeconomics. *Journal of Economic Surveys*, 31(5):1152–1182.
- Furnham, A. and Wilson, E. (2011). Gender differences in estimated salaries: A UK study. *Journal of Socio-Economics*, 40(5):623–630.
- García-González, J., Forcén, P., and Jimenez-Sanchez, M. (2019). Men and women differ in their perception of gender bias in research institutions. *PLOS ONE*, 14(e0225763).
- Gräbner, C. (2018). How to relate models to reality? An epistemological framework for the validation and verification of computational models. *Journal of Artificial Societies and Social Simulation*, 21(3).
- Grüne-Yanoff, T. (2009). Learning from minimal economic models. *Erkenntnis*, 70(1):81–99.
- Grüne-Yanoff, T. and Verreault-Julien, P. (2021). How-possibly explanations in economics: anything goes? *Journal of Economic Methodology*, 28(1):114–123.
- Haaland, I. and Roth, C. (2021). Beliefs about racial discrimination and support for pro-black policies. *Review of Economics and Statistics*, forthcoming.
- Hamilton, M. J., Milne, B. T., Walker, R. S., Burger, O., and Brown, J. H. (2007). The complex structure of hunter–gatherer social networks. *Proceedings of the Royal Society B: Biological Sciences*, 274(1622):2195–2203.
- Hampton, M. B. and Heywood, J. S. (1993). Do workers accurately perceive gender wage discrimination? *ILR Review*, 47(1):36–49.
- Harting, P. and Radi, D. (2020). Residential segregation: The role of inequality and housing subsidies. *Journal of Economic Behavior & Organization*, 178:801–819.
- Hauser, O. P. and Norton, M. I. (2017). (Mis)perceptions of inequality. *Current Opinion in Psychology*, 18:21–25.

- Hu, X. and Liang, C.-Y. (2022). Does income redistribution prevent residential segregation? *Journal of Economic Behavior & Organization*, 193:519–542.
- Jagielski, M. and Kutner, R. (2013). Modelling of income distribution in the European Union with the fokker–planck equation. *Physica A: Statistical Mechanics and its Applications*, 392(9):2130–2138.
- Karlsson, N., Gärling, T., Dellgran, P., and Klingander, B. (2005). Social comparison and consumer behavior: When feeling richer or poorer than others is more important than being so. *Journal of Applied Social Psychology*, 35(6):1206–1222.
- Kets, W. and Sandroni, A. (2019). A belief-based theory of homophily. *Games and Economic Behavior*, 115:410–435.
- Kinder, D. R. and Winter, N. (2001). Exploring the racial divide: Blacks, whites, and opinion on national policy. *American Journal of Political Science*, pages 439–456.
- Kraus, M. W., Onyeador, I. N., Daumeyer, N. M., Rucker, J. M., and Richeson, J. A. (2019). The misperception of racial economic inequality. *Perspectives on Psychological Science*, 14(6):899–921.
- Kraus, M. W., Rucker, J. M., and Richeson, J. A. (2017). Americans misperceive racial economic equality. *Proceedings of the National Academy of Sciences*, 114(39):10324–10331.
- Krieger, N. (1992). Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *American Journal of Public Health*, 82(5):703–710.
- Kunze, A. (2018). The gender wage gap in developed countries. In Averett, S. L., Argys, L. M., and Hoffman, S. D., editors, *Oxford Handbook of Women and the Economy*, pages 369–394. Oxford University Press, Amsterdam.
- Larson, J. M. (2017). The weakness of weak ties for novel information diffusion. *Applied Network Science*, 2(1).
- Leo, Y., Fleury, E., Alvarez-Hamelin, J. I., Sarraute, C., and Karsai, M. (2016). Socioeconomic correlations and stratification in social-communication networks. *Journal of The Royal Society Interface*, 13(20160598).
- Linos, K. and West, M. (2003). Self-interest, social beliefs, and attitudes to redistribution. Re-addressing the issue of cross-national variation. *European Sociological Review*, 19(4):393–409.
- Mac Carron, P., Kaski, K., and Dunbar, R. (2016). Calling Dunbar’s numbers. *Social Networks*, 47:151 – 155.
- Mäki, U. (2009). MISSing the world. Models as isolations and credible surrogate systems. *Erkenntnis*, 70:29–43.
- Malul, M. (2021). (Mis)perceptions about the gender gap in the labor market. *Forum for Social Economics*, forthcoming.

- Manski, C. (1981). *Structural analysis of discrete data with econometric applications*. MIT Press, Cambridge, Mass.
- Mayerhoffer, D. M. and Schulz, J. (2022). Marginalisation and misperception: Perceiving gender and racial wage gaps in ego networks. In Benito, R. M., Cherifi, C., Cherifi, H., Moro, E., Rocha, L. M., and Sales-Pardo, M., editors, *Complex Networks & Their Applications X*, pages 768–779, Cham. Springer International Publishing.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.
- Milli, L. (2021). Opinion dynamic modeling of news perception. *Applied Network Science*, 6(1).
- Morgan, J. and Kelly, N. J. (2017). Social patterns of inequality, partisan competition, and Latin American support for redistribution. *Journal of Politics*, 79(1):193–209.
- Newman, M. E. (2009). Random graphs with clustering. *Physical Review Letters*, 103(5):058701.
- Newman, M. E., Strogatz, S. H., and Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(2):026118.
- Nirei, M. and Souma, W. (2007). A two factor model of income distribution dynamics. *Review of Income and Wealth*, 53(3):440–459.
- Oancea, B., Andrei, T., and Pirjol, D. (2017). Income inequality in Romania: The exponential-Pareto distribution. *Physica A: Statistical Mechanics and its Applications*, 469:486–498.
- Oprea, R. (2020). What makes a rule complex? *American Economic Review*, 110(12):3913–3951.
- Pfeifer, C. and Stephan, G. (2019). Why women do not ask: gender differences in fairness perceptions of own wages and subsequent wage growth. *Cambridge Journal of Economics*, 43(2):295–310.
- Quadagno, J. and Blekesaune, M. (2003). Public attitudes towards welfare state policies. *European Sociological Review*, 19(5).
- Schneider, M. (2015). Revisiting the thermal and superthermal two-class distribution of incomes. *European Physical Journal B*, 88(1):1–10.
- Schulz, J., Mayerhoffer, D. M., and Gebhard, A. (2022). A network-based explanation of inequality perceptions. *Social Networks*, 70:306 – 324.
- Settele, S. (2022). How do beliefs about the gender wage gap affect the demand for public policy? *American Economic Journal: Economic Policy*, 14(2):475–508.
- Shaikh, A. (2017). Income distribution, econophysics and Piketty. *Review of Political Economy*, 29(1):18–29.
- Shaikh, A. and Jacobo, J. E. (2020). Economic arbitrage and the econophysics of income inequality. *Review of Behavioral Economics*, 7:299–315.

- Shaikh, A., Papanikolaou, N., and Wiener, N. (2014). Race, gender and the econophysics of income distribution in the USA. *Physica A: Statistical Mechanics and its Applications*, 415:54–60.
- Silva, A. C. and Yakovenko, V. M. (2004). Temporal evolution of the “thermal” and “superthermal” income classes in the USA during 1983–2001. *Europhysics Letters (EPL)*, 69(2):304.
- Smith, J. A., McPherson, M., and Smith-Lovin, L. (2014). Social distance in the United States. *American Sociological Review*, 79(3):432–456.
- Stelzner, M. and Bahn, K. (2022). Discrimination and monopsony power. *Review of Black Political Economy*, 49(2):152–174.
- Talaga, S. and Nowak, A. (2020). Homophily as a process generating social networks: Insights from social distance attachment model. *Journal of Artificial Societies and Social Simulation*, 23(2):6.
- Tao, Y., Wu, X., Zhou, T., Yan, W., Huang, Y., Yu, H., Mondal, B., and Yakovenko, V. M. (2019). Exponential structure of income inequality: Evidence from 67 countries. *Journal of Economic Interaction and Coordination*, 14(2):345–376.
- Törnqvist, L., Vartia, P., and Vartia, Y. O. (1985). How should relative changes be measured? *The American Statistician*, 39(1):43–46.
- Virtanen, M., Kivimäki, M., Kouvonen, A., Elovainio, M., Linna, A., Oksanen, T., and Vahtera, J. (2007). Average household income, crime, and smoking behaviour in a local area: The Finnish 10-town study. *Social Science & Medicine*, 64(9):1904–1913.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442.
- Williams, M. J., Elizabeth, L. P., and Spencer-Rodgers, J. (2010). The masculinity of money: Automatic stereotypes predict gender differences in estimated salaries. *Psychology of Women Quarterly*, 34(1):7–20.
- Wu, K. (2021). Invisibility of social privilege to those who have it. *Academy of Management Proceedings*, 2021(1).
- Xu, Y., Santi, P., and Ratti, C. (2021). Beyond distance decay: Discover homophily in spatially embedded social networks. *Annals of the American Association of Geographers*, 112(2):505–521.
- Zhou, W.-X., Sornette, D., Hill, R. A., and Dunbar, R. I. (2005). Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B: Biological Sciences*, 272(1561):439–444.

# A Network Approach to Consumption

---

Joint work with D. M. Mayerhoffer.

Revised version submitted with D. M. Mayerhoffer as “A Network Approach to Expenditure Cascades” to the Review of Behavioral Economics.

We would like to express our gratitude to Yunus Aksoy, Jan Behringer, Domenico Delli Gatti, Anna Gebhard, Gabriel Lozada, Özlem Onaran, Benjamin Tippet, Jan Weber, Rafael Wildauer and the participants of the research colloquium at the University of Utah as well as the 12th Annual PhD Student Conference of the Post-Keynesian Economics Society, especially to our designated discussant Caverzasi Eugenio, the participants of the 25th conference of the Forum for Macroeconomics and Macroeconomic Policies in Berlin, of the First Philosophy and Economics Conference in Vienna, of the Workshop on Economics with Heterogeneous Interacting Agents 2021, of the Networks in a Globalized Work Conference 2022 and of the Annual Conference of the European Association for Evolutionary Political Economy in Naples 2022 for many helpful comments and suggestions that greatly improved our paper and sharpened our focus. Finally, we also thank Marius Theiler and Carsten Källner for their able research assistance. All remaining errors are, of course, ours.

---

## Abstract

The nexus between debt and inequality has attracted considerable scholarly attention in the wake of the global financial crisis. One prominent candidate to explain the striking co-evolution of income inequality and private debt in this period has been the theory of upward-looking consumption externalities leading to expenditure cascades. We propose a parsimonious model of upward-looking consumption at the micro level mediated by perception networks with empirically plausible topologies. This allows us to make sense of the ambiguous empirical literature on the relevance of this channel. Up to our knowledge, our approach is the first to make the reference group to which conspicuous consumption relates explicit. Our model, based purely on *current* income, replicates the major stylised facts regarding micro consumption behaviour and is thus observationally equivalent to the workhorse permanent income hypothesis, without facing its dual problem of ‘excess smoothness’ and ‘excess sensitivity’. We also demonstrate that the network topology and segregation has a significant effect on consumption patterns which has so far been neglected.

---

---

## Contents of this Paper

7.1	Introduction . . . . .	199
7.2	Related Literature . . . . .	201
7.3	Model . . . . .	204
7.3.1	Network Generation . . . . .	204
7.3.2	Individual Perception and Consumption . . . . .	207
7.4	Results . . . . .	210
7.4.1	Micro Level Patterns . . . . .	211
7.4.2	Inequality and Aggregate Savings . . . . .	215
7.5	Discussion . . . . .	218
	References . . . . .	221

---

## 7.1 Introduction

When the potentially destabilising effects of both economic inequality and private debt have become apparent in the wake of the global financial crisis, the nexus between the two has also increasingly attracted scholarly interest. One prominent candidate to explain the striking co-evolution of inequality and private debt in this period has been the theory of upward-looking consumption externalities (Van Treeck 2014). Thereby, the increases in the top relative incomes are assumed to have triggered debt-financed conspicuous consumption of the poorer segment of the population. While several microeconomic studies testify to the relevance of such conspicuous consumption (Heffetz 2011; Bertrand and Morse 2016; Bricker et al. 2020), the macroeconomic evidence on this channel has been somewhat limited and ambiguous (Wildauer and Stockhammer 2018; Bofinger and Scheuermeyer 2019). We propose a parsimonious model of upward-looking consumption mediated by perception networks with empirically plausible topologies to reconcile these two contradictory findings. Up to our knowledge, our approach is the first to make the reference group to which conspicuous consumption relates explicit.<sup>1</sup> As predicted by the (upward-looking) relative income hypothesis, aggregate savings unanimously decrease in inequality. However, endogenously evolving perception networks featuring homophily might mitigate this effect because an increase in inequality also increases segregation within the network. For high degrees of homophily, the effect vanishes, potentially explaining the mismatch between micro-behaviour and macro-aggregates.

Our model is an attempt to answer the challenge posed by Nelson and Consoli (2010) that its Schumpeterian heritage lets evolutionary economics exhibit a supply-side bias and neglect demand-side factors. In particular, we build on two central themes of evolutionary economics: The relevance of imitation and of meso-level structures for macroeconomic outcomes. Following Veblen (1899), individuals engage in conspicuous consumption activity, triggering imitative and emulative behaviour by individuals with lower social status (Nelson and Consoli 2010). Recent research indeed suggests conspicuous consumption itself to be likely of evolutionary origin, though emerging not from sexual selection but rather cultural evolution within contemporary market societies (Koliofotis 2022). This culturally evolutionist perspective implies that conspicuous consumption is a socially learned activity acquired through observation and imitation (Witt 2001; Koliofotis 2022). As a direct corollary, individuals must be able to *observe* others' consumption practices in order to emulate them, which we formalise by a plausible perception network following Schulz et al. (2022). This complex interaction at the meso-level thus mediates and shapes

---

<sup>1</sup>The foundational contribution by Frank et al. (2014) only builds on classifications according to income deciles and does not include a discussion of inner-decile links that Schulz et al. (2022) find to be of crucial importance for perceptions.

the effect of micro-level behaviour (i.e., individual consumption) on macro-level outcomes (i.e., expenditure distributions and aggregate consumption levels) (Foster 2021).

The argument that intersubjective comparisons can be powerful personal motivators traces back to at least Adam Smith's (1776) *Wealth of Nations*.<sup>2</sup> He famously argued that even something as simple as wearing a linen shirt might be a "necessary of life in England" (Smith 1776, vol. ii, p. 352) at the time of his writing to signal that the wearer conforms to custom. Status signalling has been a central tenet, especially of heterodox theories of consumption behaviour ever since (Veblen 1899; Duesenberry 1949; Witt 2001; Frank et al. 2014). However, while most those theories focus on luxury goods, Smith (1776) reminds us that status goods are essentially defined by custom, and that their signalling is not an inherent feature of them. Therefore, status consumption is not confined to a specific social class that can afford expensive luxury goods but is a human tendency that might apply to all individuals across the whole income distribution. Yet, not all consumption is status consumption: Keynes (1972) divided goods into two types: Non-positional goods relating to material needs and status goods relating to needs "which are relative in the sense that we feel them only if their satisfaction lifts us above, makes us superior to, our fellows" (Keynes 1972, p. 326). In the terminology of an evolutionary theory of consumption, this taxonomy would roughly correspond to basic, innate needs 'needs' and socially constructed 'wants' (Witt 2001; Rengs and Scholz-Wäckerle 2019). Psychological evidence indeed corroborates the conjecture that status consumption is upward-looking (Frank 1985) and that both absolute and relative motives matter (Alpizar et al. 2005).

Our model includes both the Smithian and Keynesian insight and lets *all* consumers have a propensity for *upward-looking* status consumption, with the whole population consuming a basket of both status and non-status goods. This parsimonious consumption rule based only on *current* income replicates all the major stylised facts regarding micro consumption behaviour. It is thus at least observationally equivalent to the workhorse permanent income hypothesis, without facing its problems of 'excess smoothness' and 'excess sensitivity' (Meghir 2004).

Despite their emphasis on signalling and perception, formal theories of conspicuous consumption seldomly explicate to which others the consumers exactly aspire to catch up to. Typically, only global or class average consumption levels are considered (Alpizar et al. 2005; Rengs and Scholz-Wäckerle 2019; Alvarez-Cuadrado and Van Long 2011; Petach and Tavani 2021), meaning that local structures are not part of these explanations. Contrary to that, the sociological literature indicates that empirical social networks are far from random and consequently that human interaction is strongly structured. The major contribution of this paper is to show that the impact of this structure, most notably homophily or the tendency to link to individuals that are similar to oneself (Lazarsfeld and Merton 1954), is non-negligible and might help explain some of the recent empirical

---

<sup>2</sup>Cf. Sen's (1984) foundational discussion about absolute and relative dimensions of poverty.



puzzles. The most notable of these puzzles is the mismatch between significant findings in the applied micro literature and the ambiguous findings from more macroeconomically oriented approaches on positional concerns. We propose a new channel, where income inequality not only incentivises conspicuous consumption but also increases segregation due to income homophily, thereby decreasing *perceived* inequality and mitigating the initial effect of *actual* inequality. Bertrand and Morse (2016) document the relevance of this channel for residential segregation, and we hope to inspire further empirical studies with our model exercise, especially in the direction of occupational or educational segregation in relation to consumption decisions.

The remainder of this paper is organised as follows: Section 2 situates our model within the pertinent literature and shows how both sociological and economic theory inform our modelling choices. Section 3 introduces our formal model both regarding consumption behaviour and the underlying social network. Section 4 presents our simulation results. Within the section, we demonstrate both that the emergent expenditure distributions are consistent with the relevant stylised facts and that the endogenous adjustment of the social networks can generate both nonlinearities and a large variety of point elasticities, underlining the potential relevance of the segregation channel. The final section 5 concludes and discusses possible avenues for further research, particularly possible empirical applications and modelling extensions to include our partial model of consumption decisions within a full-fledged macro framework.

## 7.2 Related Literature

The advent of Friedman’s (1957) permanent income hypothesis was a paradigm change in the theory of consumption behaviour and shifted the focus from explaining consumption based on current income to explaining it with the concept of permanent income composed of unobservable stochastic income shocks. The subsequent literature has focused mainly on refining Friedman’s initial hypothesis by including expectations, precautionary savings motives or liquidity constraints to bring it closer to the data (Palley 2010; Meghir 2004, for surveys). Our focus is different. We abstract from any intertemporal expectation formation or additional liquidity constraints and show how the emergent expenditure distributions relating to *current* income only can give rise to the relevant empirical stylised facts of consumption behaviour purely from human interaction within plausible social networks. Nonetheless, expectations regarding future income streams are potentially relevant for consumption behaviour, and our model merely demonstrates that the empirical findings on consumption expenditures both on the micro and on the macro level do not necessarily imply intertemporal optimisation over (largely unobservable) future income streams, as is typically argued in the literature (Battistin et al. 2009). In contrast to these demanding assumptions on (essentially unbiased or model-consistent) individual

expectations, we make the perceptions of each agent fully explicit without any kind of expectation formation on future income streams. Thereby, we bring two strands of literature together to make sense of a third one: Namely, we combine the rich extant literature on empirical social networks with the empirical evidence on upward-looking consumption externalities to provide a new explanation for the ambiguous empirical findings on the aggregate savings-inequality nexus.

The topology of empirical social networks exhibits salient and universal features that can serve as stylised facts to validate artificial graphs. These include the small-world property which holds across many different domains (Galaskiewicz 2007; Newman 2001; Uzzi and Spiro 2005; Weeden and Cornwell 2020), homophily in tie-formation, especially for economic class (Boucher 2015; Cepić and Tonković 2020; Huckfeldt 1983; Malacarne 2017; Mayer and Puller 2008) and their sparsity, especially when focusing on close ties (Mac Carron et al. 2016; De Giorgi et al. 2020). Schulz et al. (2022) present a model of homophilic tie formation that can generate these stylised facts of empirical graph topologies as well as being able to replicate all the relevant findings on positional self-assessment and perceptions of inequality. Since our argument builds on status consumption, which rests on the individuals' perception of consumption, we adopt this modelling framework as an algorithm for plausible social networks.

One can understand this as individuals trying to maximise their social capital (Coleman 1988; Annen 2003) by picking those others as social contacts who are like them in certain respects because this minimises the transaction costs of social interaction (Akerlof 1997). Such transaction costs may simply be travel time if the social contacts do not live nearby, effort necessary to understand each other's cultural and social background or their interests. However, as factors like residential area (Hu and Liang 2022; Harting and Radi 2020), education (Smith et al. 2014; Leo et al. 2016), ethnicity (Chandra 2000), lifestyle (Virtanen et al. 2007), or even health (Krieger 1992) correlate with income, the latter constitutes a decent proxy for the actual homophilic behavioural patterns. Moreover, there is also homophily in income surfacing in, e.g. mobile phone communication (Xu et al. 2021; Leo et al. 2016; Fixman et al. 2016). Therefore, while people may not willingly choose their social contacts based on income proximity, their actual choices amount to social contacts as if selected homophilic in income.

The importance of interpersonal comparisons in consumption and income is well established for the microeconomic level. Among others, Bertrand and Morse (2016), Bricker et al. (2020) and Heffetz (2011) find significant effects of interpersonal comparisons for individual consumption in the US, Jinkins (2016) for the US and China, Quintana-Domeque and Wohlfart (2016) for the UK, Drechsel-Grau and Schmid (2014) for Germany and Alpizar et al. (2005) for Costa Rica. However, the choice of reference groups within these studies often lacks granularity, underlining the need for more plausible models of interaction in networks. The empirical psychological literature reviewed in chapter 2 of Frank

(1985) only establishes that conspicuous consumption is generally *upward-looking*. Given upward-looking consumption externalities, uneven income gains that are skewed to the top can trigger “expenditure cascades” if the poorer try to emulate the consumption behaviour of the richer (Frank et al. 2014). Indeed, several studies for the US also document the relevance of inequality for private debt-buildup of the poor to enable consumption increases, also in line with the relative income hypothesis (Agarwal et al. 2020; Carr and Jayadev 2015; Christen and Morgan 2005; Schmid and Drechsel-Grau 2013).

Given the apparently unambiguous results from more micro-level studies, it seems surprising that this view has not manifested itself in a consensus on the macro-level association of inequality and (private) savings. While Koo and Song (2016) report that aggregate saving rates increase with income inequality due to the rich saving more, Klein (2015), Perugini et al. (2016), Behringer and Van Treeck (2018), Behringer and van Treeck (2019) and Petach and Tavani (2021) find evidence for a negative relationship between savings and inequality, in line with the expenditure cascades hypothesis. Most studies fail to find a significant effect in either direction, though (Bordo and Meissner 2012; Cuaresma et al. 2018; Gu and Huang 2014; Gu et al. 2015; Leigh and Posso 2009; Schmidt-Hebbel and Servén 2000; Wildauer and Stockhammer 2018). Cuaresma et al. (2018), Gu and Huang (2014) and Gu et al. (2015) highlight the potential relevance of unobservable intermittent country characteristics shaping the relationship between savings decisions and inequality, which might in part explain why Bofinger and Scheuermeyer (2019) find a strongly non-monotonic relationship between both.<sup>3</sup> Furthermore, Wildauer and Stockhammer (2018) highlight the time-scale of adjustments and show that with the proper controls and examining a long-run relationship, the effect of inequality on savings seems to be significant.

We conclude that a proper model of expenditure cascades has to make sense of the variability of effect size as a function of the considered time scale. Moreover, such a model should give a plausible account of the origin of cross-country heterogeneity. To demonstrate the relevance of network structures for these phenomena, we use an otherwise purely upward-looking consumption rule inspired by Frank et al. (2014) that should in and of itself give rise to an unambiguously negative relation between savings and inequality but can generate a large variety of effect sizes as a function of network topology. Finally, we require our model to be consistent with the major stylised facts on empirical expenditure distributions for validation at the micro-level.

The empirical literature has identified four stylised facts regarding consumption expenditures and their relation to (current) incomes: (i) individual average propensities to consume (APCs) tend to decrease in (current) income (Dynan et al. 2004; Clementi and

---

<sup>3</sup>Bofinger and Scheuermeyer (2019) include country- and time-fixed effects into their estimation. These might, however, be insufficient to capture time-variant country idiosyncrasies, as our endogenously evolving social network might suggest.

Gianmoena 2017); (ii) population-level APCs remain roughly constant with respect to changes in total income (Kuznets 1942); (iii) the distribution of consumption expenditures is more homogeneous than the distribution of current income (Krueger and Perri 2006; Jappelli and Pistaferri 2010) and (iv) the distribution of consumption expenditures is (at least as a first-order approximation) well fit by a log-normal distribution (Battistin et al. 2009; Brzozowski et al. 2010; Chakrabarti et al. 2018; Hohnisch et al. 2002; Fagiolo et al. 2010; Ghosh et al. 2011; Toda 2017), while the distribution of current income is not (Drăgulescu and Yakovenko 2001; Silva and Yakovenko 2004; Tao et al. 2019). The last stylised fact constitutes evidence for the log-normality of expenditure distributions since several of those studies also find significant deviations from the log-normal benchmark, particularly for their upper tails (Chakrabarti et al. 2018; Fagiolo et al. 2010; Ghosh et al. 2011; Toda 2017). In light of these findings, we also discuss the relevant parameter range of our model for which log-normality holds, namely, for intermediate ranges of social consumption to make sense of possible deviations from the log-normal benchmark that can also emerge within our model.

## 7.3 Model

This section gives a content-oriented presentation; we provide a technical description following the ODD protocol, the NetLogo implementation of the simulation model and our raw data upon request. The model consists of two distinct phases running in sequential order, where phase one builds the environment and determines available information for the consumption procedure in the second phase:

1. Network generation
  - Agent initialisation and income allocation
  - Homophilic linkage
2. Individual perception and consumption

### 7.3.1 Network Generation

There are 1000 agents in the model; each agent draws their income from an exponential distribution with a mean of  $\lambda = 1$ . Such a distribution replicates the empirical observed (pre-tax or market) income distributions in various industrialised countries for the vast majority of individuals (Drăgulescu and Yakovenko 2001; Silva and Yakovenko 2004; Dos Santos 2017; Tao et al. 2019).<sup>4</sup> Thus, one can understand the model population

---

<sup>4</sup>Interestingly, Dos Santos (2017) shows how an exponential wage distribution emerges from social comparisons, much like the consumption externality we model in our consumption rule. It, therefore, appears promising to bring these two modelling approaches together.

as constituting a representative sample of empirical populations of these countries. The upper tail of 1 to 5 % of the income distributions empirically follows a Pareto law (Silva and Yakovenko 2004). We deliberately choose to exclude this small minority from our model since their population size would induce another degree of freedom in our model, and we want to demonstrate that segregation is indeed endogenous and not driven by differences in the actual income regime. The theoretical Gini coefficient of any exponential distribution is fixed at 0.5. To understand the relationship between differences in actual inequality levels and consumption, we therefore use log-normal distributions for the distribution of income. While such log-normal distributions do not represent empirical findings, they allow for varying the inequality level through their dispersion parameter  $\sigma$  without changing the overall distribution shape. Hence, the present model employs the log-normal distribution solely to evaluate the relationship between actual income inequality, individual perceptions and consumption behaviour. Agents store their true income decile for evaluation purposes, too.

Each agent draws five other agents to link to. Like for real-world networks, links are therefore created *by agents*, not imposed on them. The number of five-link choices represents the closest layer of intense contacts identified in empirical studies (Zhou et al. 2005; Hamilton et al. 2007; Mac Carron et al. 2016)<sup>5</sup>. While this closest layer consists of differently related people for different individuals (Karlsson et al. 2005), Wilcox and Stephen (2013) empirically identify this layer as the relevant reference group for social consumption. Moreover, we have carried out sensitivity analyses and found a larger number of links to have little impact on overall simulation results. Agent  $j$ 's weight in agent  $i$ 's draw is denoted by  $\Omega_{ij}$  and determined as follows:

$$\Omega_{ij} = \frac{1}{\exp[\rho |Y(j) - Y(i)|]} \quad (7.1)$$

The relative weights in the draws are a function of the homophily strength and the respective income levels:  $Y$  denotes the income of an agent, and  $\rho \in \mathbb{R}_0^+$  denotes the homophily strength in income selection, externally set, and identical for all agents.  $\rho = 0$  represents a random graph, and for an increasing positive value of  $\rho$ , an agent becomes ever more likely to pick link-neighbours with incomes closer to their own. The link function's exponential character ensures that those with large income differences become unlikely picks even at low homophily strengths. For an extensive analytical discussion of this linkage behaviour, see Schulz et al. (2022). While our linkage function represents a tractable reduced-form representation of the empirically well-established homophily, Falk and Knell (2004) demonstrate that endogenous reference groups can emerge from the social comparisons of agents aiming to self-improve and self-enhance. Similarly, our

---

<sup>5</sup>For a recent review on the large literature on 'Dunbar's number', cf. the first section of Mac Carron et al. (2016).

weighting function can be interpreted as agents selecting contacts as if minimising the income distance that they observe as a noisy signal. As discussed in Section 7.2, this assumption does not require individuals to consciously form their social ties based on income proximity. Instead, they may exhibit homophily in e.g. education level, ethnicity, or lifestyle; but thereby the individuals act as if there was explicit income homophily because the aforementioned factors correlate with income. As a consequence, segregation increases with income inequality, as is well documented in the literal sense of residential segregation (Wheeler and La Jeunesse 2008; Watson 2009; Reardon and Bischoff 2011; Chen et al. 2012; Tóth et al. 2021). However, linkages in our sense extend beyond this literal geographic sense and essentially imply that consumption decisions are observable, be it due to common workplaces, family ties or any other form of linkage.

Formally, we assume that the utility of agent  $i$  connecting to agent  $j$  follows an additively separable utility function with agents receiving disutility proportionally to the absolute income distance to the income of agent  $j$  and a random error term, i.e.,

$$U_{ij} = -\rho|Y(i) - Y(j)| + \epsilon_{ij}. \quad (7.2)$$

When the distribution of  $\epsilon_{ij}$  is identically and independently distributed according to an extreme value distribution, we can express the choice probability of agent  $i$  as

$$p_{ij} = \frac{\exp[-\rho \cdot |Y(j) - Y(i)|]}{\sum_{k \in M \setminus i} \exp[-\rho \cdot |Y(k) - Y(i)|]}, \quad (7.3)$$

with  $M \setminus i$  as the set of all agents except  $i$  with size  $N - 1$  which is equivalent to the weights in eq. (7.1) translated into probabilities (Hoffman and Duncan 1988). Our weighting function is thus an application of the discrete choice approach developed and popularised by Manski and McFadden (1981). The random utility model appears to us to be particularly appropriate here, since income is a rather salient characteristic to determine a good ‘fit’ within a social network but of course might depend on other characteristics that are not directly observable and thus modelled to be stochastic.<sup>6</sup> The homophily parameter  $\rho \in (0, \infty)$  is then simply the intensity of choice parameter with  $\rho \rightarrow \infty$  implying that  $i$  chooses agent  $j$  to link to with certainty who has minimal income distance, i.e.,  $p_{ij} \rightarrow 1$ . In this sense, the weighting function in eq. (7.1) is plausibly microfounded and can now be considered the workhorse choice rule in behavioural macroeconomics (Franke and Westerhoff 2017). Franke and Westerhoff (2017) also survey evidence from several lab

---

<sup>6</sup>Note, however, that this choice rule implies the axiom of Independence of Irrelevant Alternatives (IIA) to hold for all alternatives and was in fact explicitly designed to do so (Manski and McFadden 1981), i.e., the probability of choosing between  $j$  and  $k$  to be independent of the probability of choosing  $l$  (Luce 1977). IIA might be considered a good first-order approximation to model homophilic choice, but especially within social networks, knowing one agent  $j$  might indeed increase the likelihood of knowing another agent  $l$  that is acquainted with  $j$ . It might thus prove interesting to extend and generalise the above rule for correlated choice to examine the effects on the network topology in further research.

experiments in different macroeconomic contexts that discrete choice is indeed consistent with the data, while Anufriev and Hommes (2012) and Anufriev et al. (2018) provide laboratory evidence for the discrete choice approach for financial markets.

The emergent networks from this graph-formation process generate networks that replicate the relevant stylised facts regarding both overall inequality perceptions (Schulz et al. 2022) as well as perceptions of intergroup inequalities, i.e., the gender and racial wage gaps (Mayerhoffer and Schulz 2022b,a). Since perceptions are empirically based on these social samples, the validity in terms of perceptions these studies establish translates into external validity of the network we employ to model everyday perceptions. The fact that the generated networks also exhibit real-world topological properties of social networks adds to the plausibility of the generating process (Schulz et al. 2022).

### 7.3.2 Individual Perception and Consumption

Generally, positional concerns can be modelled both within an intertemporal maximisation framework with explicit expectation formation (Drechsel-Grau and Schmid 2014) or building merely on current income without any expectations relating to future income streams (Frank et al. 2014). We choose the latter option for two reasons: Firstly, we aim to keep assumptions on individual behaviour as minimal as possible to preserve maximal generalisability. The workhorse consumption model built on Euler equations as famously introduced by Hall (1978) builds on rational expectations that are very demanding as an assumption. Empirically, Pesaran and Weale (2006) find little evidence for the rational expectations hypothesis. Theoretically, Hendry and Mizon (2014) show that a violation of the usual regularity conditions in stochastic processes like non-stationarity or structural breaks implies that rational forecasts conditioned on past performance are suboptimal, essentially since the application of the law of iterated expectations is hindered. Thus, it seems highly unlikely that the rational expectations benchmark indeed holds in the real world, especially if income processes are very noisy. Secondly, the dual problem of ‘excess smoothness’ and ‘excess sensitivity’ for calibrated consumption models of permanent income hints at the shortcomings of the permanent income hypothesis (Meghir 2004), that is, consumption reacting too strongly (weakly) to (un)anticipated income shocks.<sup>7</sup> By contrast, our model is based explicitly on current income (and intersubjective comparisons) and does therefore not run into the problem that consumption is tracking current incomes too closely. We also do not need to impose an unrealistic degree of heterogeneity or additional assumptions to match the empirically observed heterogeneity in consumption propensities (Jappelli and Pistaferri 2014).

---

<sup>7</sup>The commonly employed habit persistence extension to remedy these problems yields coefficients of habit formation that are typically deemed much too high to be plausible (Pesaran 2015, p. 887 – 892, and the references therein, for a more detailed discussion). Coincidentally, Binder and Pesaran (2001) demonstrate that habit persistence in conjunction with social comparisons might be able to solve both problems simultaneously.

In particular, the only heterogeneity we impose is the (pre-validated) income distribution. Agents consume maximising utility  $U(\cdot)$  with Cobb-Douglas preferences and derive (linear) disutility the stronger their consumption in isolation given by  $\tilde{C}(i)$  falls short of the highest consumption level  $C(j|i)$  they observe within their ego network with intensity parameter  $c > 0$ . With this assumption, we follow the tradition of additive comparison utility in contrast to ratio comparison utility, where others' consumption enters as a ratio to own isolated consumption (Alpizar et al. 2005).<sup>8</sup> Apart from that the utility function is rather standard with agents deriving utility both from consumption with elasticity parameter  $w$  and saving with elasticity parameter  $(1 - w)$  to capture intertemporal motives in the most parsimonious fashion. The choice variable is the average propensity to consume  $\gamma$ , i.e., the fraction of current income  $Y$  an agent consumes. The utility function of agent  $i$  thus reads

$$U(\gamma_i; Y(i), C(j|i), \tilde{C}(i), c, w) = (\gamma_i Y(i) - c \cdot (C(j|i) - \tilde{C}(i)))^w ((1 - \gamma_i) Y(i))^{1-w} \quad (7.4)$$

The utility function for consumption in isolation,  $\tilde{U}$  with choice variable  $\tilde{\gamma}_i$  is given by eq. (7.4) for  $c = 0$  and thus is given by

$$\tilde{U}_i(\tilde{\gamma}_i; w, Y(i)) = (\tilde{\gamma}_i Y(i))^w (Y(i) \cdot (1 - \tilde{\gamma}_i))^{1-w}. \quad (7.5)$$

The FOC for optimal  $\tilde{\gamma}$  reads

$$\tilde{\gamma}_i = w, \quad (7.6)$$

and thus optimal consumption without social interaction is given by

$$\tilde{C}(i) = wY(i), \quad (7.7)$$

i.e., the canonical result that consumption is a constant fraction of income. Substituting  $\tilde{C}(i)$  from eq. (7.7) into eq. (7.4) yields

$$U(\gamma_i; Y(i), C(j|i), c, w) = (\gamma_i Y(i) - c \cdot (C(j|i) - wY(i)))^w ((1 - \gamma_i) Y(i))^{1-w}. \quad (7.8)$$

By the FOC we derive the optimal  $\gamma_i$  as

$$\gamma_i = \frac{wY(i) + (1 - w)c(C(j|i) - wY(i))}{Y(i)}, \quad (7.9)$$

---

<sup>8</sup>A prominent example of the additive approach is Akerlof (1997), while one example of ratio comparisons can be found in Carroll et al. (1997). Cf. also the literature review in Alpizar et al. (2005). The qualitative intuition should be equivalent in both cases, though.



and consequently consumption by noting that  $C(i) = \gamma_i Y(i)$  as

$$C(i) = wY(i) + (1 - w) \cdot c(C(j|i) - wY(i)). \quad (7.10)$$

Since we assume  $w$  and  $c$  to be constant for all agents, the consumption rule is entirely equivalent for all agents. This implies that the consumption of any agent  $i$  derived here from a utility maximising framework is a weighted sum of idiosyncratic and socially determined consumption with weights  $w$  and  $(1 - w)$ .<sup>9</sup>

We want to emphasise that the derivation by utility maximisation is not the only conceivable way to arrive at such a consumption rule. Equivalently, the rule can also be interpreted as resulting from behavioural heuristics, where  $w$  is the marginal propensity to consume out of current income and  $c$  is the propensity to consume out of the consumption differential to the individual indexed  $j$  without any reference to utility functions. In the terminology of evolutionary theories of consumption, the first term in (7.10) would capture the ‘needs’ of consumers with the latter term corresponding to the socially constructed ‘wants’ (Witt 2001; Rengs and Scholz-Wäckerle 2019). In (Post-)Keynesian jargon, these terms would instead be interpretable as capturing the intuition of the absolute and relative income hypothesis (Palley 2010), respectively. The formal representation is in this sense neutral with respect to the subtleties of marginalist, evolutionary or Post-Keynesian interpretations and can thus be incorporated in those different frameworks without much difficulty.<sup>10</sup>

Overall, intertemporal motives only feature within the model insofar as agents derive utility directly from saving in anticipation of future consumption and hence leave no role regarding expectation formation regarding potential future income gains or losses. However, our approach is also consistent with the permanent income hypothesis’s central intuition: One can also understand the interpersonal consumption comparison presented here as a proxy for intertemporal ones. Individuals might look at others they perceive to lead overall parallel lives as themselves and take the observed consumption as a predictor for what is possible for themselves. In that way, our model explains why life-cycle models replicate the four stylised facts but necessitate implausible behavioural assumptions to do so: Individuals act as if they smoothed over future incomes while actually observing some of these potential future incomes in others. Arguably, merely observing consumption de-

---

<sup>9</sup>Note that this formulation is reminiscent of the consumption rule employed in the foundational paper by Frank et al. (2014) who just postulate and not derive the consumption function from utility maximisation, though.

<sup>10</sup>In fact, even though social interdependencies seem now foreign to orthodox marginalist approaches, they were subject to large debates in the early days of marginalist consumer theory at the turn of the twentieth century (Bianchi and Sanfilippo 2015). In this sense, a heuristics-based approach might not be that far from a marginalist approach based on purposeful optimisation, with the two approaches not differing too much in the predicted consumption behaviour but only in the explanation of this behaviour.

cisions rather than forming explicit (and model-consistent) expectations over a stochastic income innovation process appears to be much more plausible.<sup>11</sup>

Given the model mechanism, the income levels  $Y(j|i)$  on which the consumption of  $i$  (directly or indirectly) depends can be indexed by an ordering  $\mathcal{R}_i = 0, 1, \dots, d_i$  where  $d_i$  is the distance of the most distant income to  $i$  on which  $C(i)$  depends. It follows that  $Y(0|i) = Y(i)$ , as the distance is then 0. By recursive substitution, we can rewrite  $C(i)$  as

$$C(i) = \sum_{j=0}^{d_i} ((1-w) \cdot c)^j \cdot (1 - (1-w) \cdot c) \cdot w \cdot Y(j|i) = \sum_{j=0}^{d_i} \beta_j Y(j|i). \quad (7.11)$$

The consumption of any individual  $i$  is thus directly dependent on their own income but depends also on all other incomes to which they are (directly or indirectly) connected. Since  $((1-w) \cdot c)$  is, per assumption, strictly below unity, the weight of a given distant income level  $((1-w) \cdot c)^j$  decays (exponentially) in the graph distance  $j = 0, 1, \dots, d_i$  to the relevant agent. While more distant incomes are thus generally higher due to social comparisons being upward-looking, they are also given less weight. This implies that  $C(i)$  is a linear combination of an exponentially distributed random variable  $Y$  with varying weights and number of terms and thus, a so-called hypoexponential mixture (Li and Li 2019; Yanev 2020). As we discuss in more detail in section 4, this directly implies two of the four stylised facts, namely, the approximate log-normality of expenditure distributions and the fact that they are robustly more homogeneous than income distributions.

Our homophilic graph formation process in conjunction with the upward-looking consumption rule thus formalises a central hypothesis by Thorstein Veblen explicitly at the micro level who conjectured that "[...] our standard of decency in expenditures, [...] is set by the usage of those next above us in reputability; until, in this way, especially in any community where class distinctions are somewhat vague, all canons of reputability and decency, and all standards of consumption, are traced back by insensible gradations to the usages and habits of thoughts of the highest social and pecuniary class" (Veblen 1899, p. 72). While the homophilic graph formation guarantees that agents only react to consumption of those that are close to them regarding their social status, eq. (7.11) shows that the consumption of the poor might indirectly depend even on the consumption of the richest agents through the network.

## 7.4 Results

The results we present are twofold: Firstly, we show that our parsimonious model setup embedded within realistic social networks can replicate the known stylised facts of empir-

<sup>11</sup>We thank Caverzasi Eugenio for pointing us to this alternative interpretation.

ical expenditure distributions at the micro-level. On the one hand, these results testify to the validity of our model assumptions. On the other hand, they also serve as proof of concept to show that a static setting may replicate empirical expenditure patterns without resorting to unobservable stochasticity, strong assumptions about credit markets and individual rationality or assuming any heterogeneity in individual consumption functions *ex-ante*. Secondly, we spell out the implications for savings and inequality at the macro level. In particular, we show that our model can generate a large variety of different elasticities of savings with respect to inequality, in line with the highly ambiguous empirical literature.

### 7.4.1 Micro Level Patterns

As laid out in section 2, we consider four stylised facts for validation with varying levels of granularity. Stylised facts (i) to (iii) describe qualitative features of expenditures that are, in principle, consistent with many different functional forms describing the distribution of expenditures. The fourth stylised fact is more stringent and imposes log-normality as the functional form. Even though log-normality is a more demanding requirement, it does not itself imply stylised facts (i) to (iii), which we, therefore, consider separately for full validation (Fagiolo et al. 2019). The first stylised fact concerns the decrease of individual APCs in income levels. Indeed, our model replicates this finding for all levels of social consumption without imposing different consumption propensities *ad hoc*. This follows directly from social consumption being purely upward-looking. If one is richer than all their link-neighbours, they do not need to catch up to anyone’s higher consumption, and hence their consumption is entirely idiosyncratic. However, social consumption accumulates as it is passed down within the network from richer to poorer individuals. Thus, poorer individuals tend to exhibit higher social consumption above their idiosyncratic consumption, increasing their personal APC. Figure 7.1 shows the APC schedules per decile for different parametrisations and demonstrates that decile APCs are indeed declining, in line with the empirical evidence. Note that this result is endogenous to our model and not imposed *ex ante*, as is frequently postulated in class-based macroeconomic models (cf. Rengs and Scholz-Wäckerle 2019, for a recent example).

As Figure 7.1 also shows, though, the precise form of this decay depends on the specific parametrisation. The  $w$  parameter determines the weight of idiosyncratic consumption. Since idiosyncratic consumption is relatively more important for the richest,  $w$  is also most relevant for the level of the total consumption of the richest decile, as is apparent from the fact that the blue and orange APC schedules with equal  $w$  but different  $c$  in Figure 7.1 almost coincide for the richest decile, whereas the red and orange schedules with equal  $c$  but different  $w$  do not. In contrast to that, the  $c$  parameter captures the strength of the ‘catching-up’ behaviour, which manifests itself in the steepness of the curve, as is

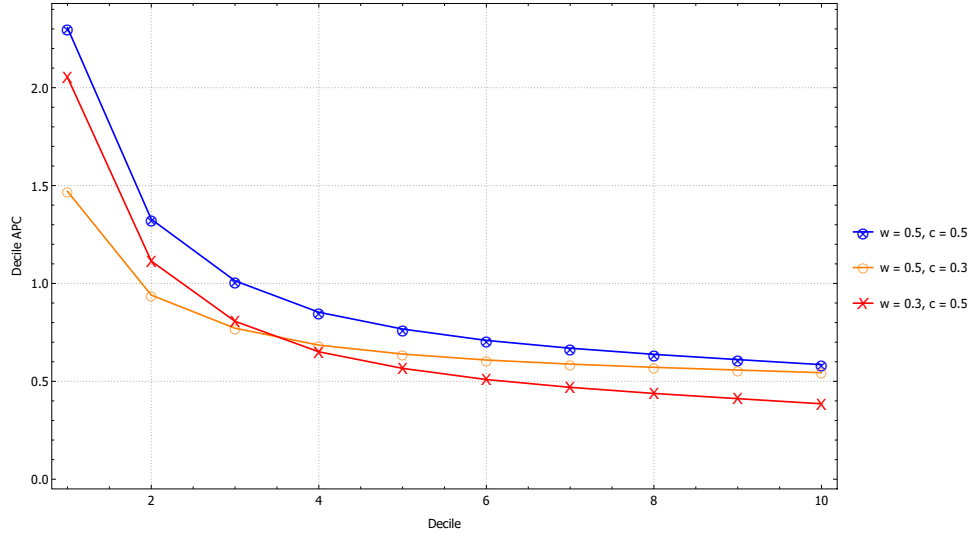


Figure 7.1: The figure plots the APCs per income decile for different parametrisations and with homophily strength  $\rho = 4$  for a single simulation run each. Lines are visual aids only.

also readily visible from comparing the orange and red APC schedules.<sup>12</sup> Regardless of the parametrisation leading to different functional forms of the APC schedules, they are, however, all declining in income decile, replicating stylised fact (i).

The most prominent fact regarding empirical expenditure distributions was the finding by Kuznets (1942) that aggregate APCs stay approximately constant for changes in aggregate income. Indeed, this was one of the major factors contributing to the paradigm change from the absolute to the permanent income hypothesis, as the absolute income hypothesis in its affine-linear variant exhibits decaying aggregate APCs in total income (Palley 2010). The invariance of aggregate APCs  $C/Y$  from changes in  $Y$  follows trivially from the recursive solution to  $C(i)$  in eq. (7.11). Assume a proportional change to  $Y$  by  $\omega \in \mathbb{R}^+$ , such that  $\tilde{Y}(i) = \omega Y(i) \forall i$  and  $\tilde{Y} = \omega Y$ . Thus,  $\omega$  is independent of  $Y(i)$  and does not change orderings  $\mathcal{R}_i$ . Consequently, since  $\omega$  cancels out, total consumption is given by the following:

$$\tilde{C} = \sum_{i=1}^n \sum_{j=0}^{d_i} \beta_j \omega Y(j|i) = \omega \sum_{i=1}^n \sum_{j=0}^{d_i} \beta_j Y(j|i) \text{ and therefore} \quad (7.12)$$

$$\frac{\tilde{C}}{\tilde{Y}} = \frac{\omega \sum_{i=1}^n \sum_{j=0}^{d_i} \beta_j Y(j|i)}{\omega \sum_{i=1}^n Y(i)} = \frac{C}{Y}. \quad (7.13)$$

<sup>12</sup>We chose these particular parametrisations, since, for  $\rho = 4$ , they appear to capture the behavior of empirical APCs reported in Clementi and Gianmoena (2017) reasonably well. Decreasing  $\rho$  tends to increase the reported effects since the income difference to the richest observed individual within a perception set tends to increase with decreasing  $\rho$  by assumption, leading to stronger ‘catching-up’. Simulation results verifying this finding are available upon request.

As a result, aggregate APCs are invariant to any positive proportional change to income, i.e., distribution-preserving changes. Whenever inequality by one of the usual measures increases, aggregate APCs increase, too, as we show in the subsection below.

As a third stylised fact, we considered that the distribution of consumption expenditures is more homogeneous than the income distribution. In our specification, this is necessarily also true for  $w < 1$ ,<sup>13</sup> since by the recursive solution in eq. (7.11), expenditure levels are a weighted sum of income levels, themselves following an exponential distribution, with varying weights and number of terms. This characteristic implies that the distribution of  $C$  is *hypoexponential*, with a coefficient of variation strictly smaller than unity, while the exponentially distributed income levels exhibit a coefficient of variation (asymptotically) equal to unity (Li and Li 2019; Yanev 2020). Intuitively, since social consumption accumulates in a cascade down the income distribution, richer individuals are relatively unaffected by status concerns, while the cumulative effect on poor individuals is much higher. Thus, increasing the level of status consumption by decreasing  $w$  or increasing  $c$  tends to equalise the expenditure distribution since the relatively higher status consumption of the poorer lets their consumption approach the expenditures of the rich that are not so much affected by status concerns.

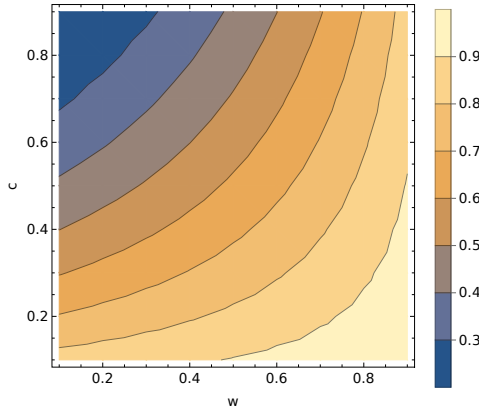


Figure 7.2: Ratio of the coefficients of variation for the expenditure and income distribution for the whole parameter space of  $w$  and  $c$  and with  $\rho = 0.5$ , averaged over 100 simulation runs.

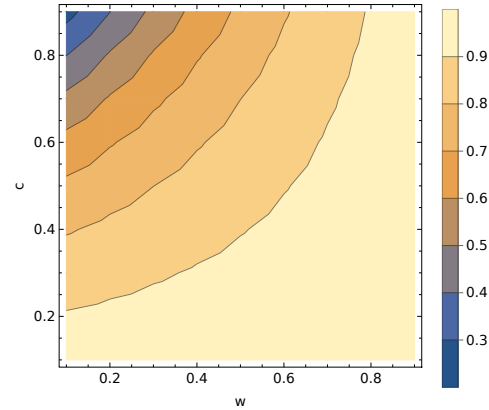


Figure 7.3: Ratio of the coefficients of variation for the expenditure and income distribution for the whole parameter space of  $w$  and  $c$  and with  $\rho = 4$ , averaged over 100 simulation runs.

This is also what we find for our numerical simulations, here for 100 runs each, as we show in Figures 7.2 and 7.3 for the parameter combinations of  $w$  and  $c$ . For all cases, the coefficient of variation is strictly below the one for the income distribution, replicating stylised fact (iii). However, the equalising tendency is much less pronounced for high

<sup>13</sup>For  $w = 1$ , we would simply recover a (rescaled) version of the exponential income distribution for expenditures, and the recursive solution in eq. (7.11) would be undefined.

levels of homophily, as can be seen in Figure 7.3 because the much more homophilic link formation tends to decrease the maximum income level each individual observes. Hence, homophily tends to hinder the transmission of social consumption.

Finally, we also find that our specification can map exponential income distributions to log-normal expenditure distributions, even though both come from completely different distributional families. The reason behind this finding is simply that the hypoexponential mixture characterising the expenditures is well approximated by a log-normal if it exhibits sufficient skewness, but the underlying rate parameters  $\lambda$  of the exponential are not too heterogeneous. A representative example is shown in Figures 7.4 and 7.5.

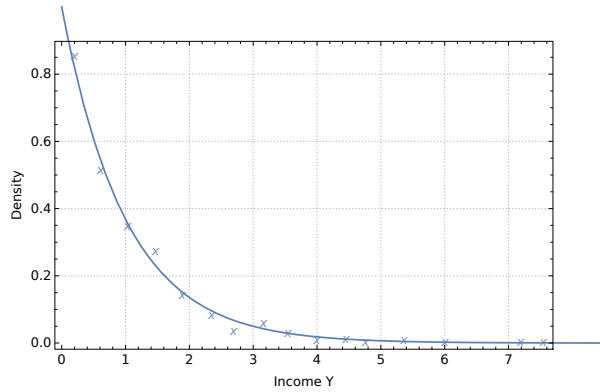


Figure 7.4: Empirical Density of Personal Incomes  $Y(i)$ .

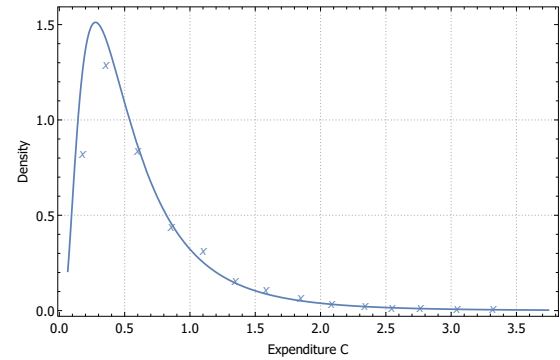


Figure 7.5: Empirical Density of Personal Expenditures  $C(i)$  for a single simulation run with  $w = 0.5$ ,  $c = 0.3$  and  $\rho = 1$ .

Per assumption, the pre-validated income distribution follows an exponential distribution with rate parameter  $\lambda = 1$ . However, the expenditure distribution is extremely well fit by a log-normal distribution, replicating stylised fact (iv). Note that this represents an entirely new generating mechanism for log-normal distributions and does not rely on any unobservable income innovation process or model-consistent expectations as are typically used in mainstream macro models. This is a relevant finding in and of itself since the log-normality of consumption expenditures is typically taken as evidence that consumption cannot be a function of *current* income and thus needs to depend on *permanent* income subject to a stochastically multiplicative growth process generating the log-normal functional form (Battistin et al. 2009). Our model exercise demonstrates that this is not necessarily true. A parsimonious model based on current income and simple behavioural rules of thumb can simultaneously replicate this stylised fact without facing the theoretical and empirical shortcomings of models in the rational expectations tradition.

Log-normality holds for large parts of the parameter space, as we show in Figures 7.6 and 7.7. The approximation only breaks down for very high levels of idiosyncratic consumption and  $w$  approaching unity, where we recover a rescaled version of the initial exponential distribution and very high levels of social consumption with low  $w$  and high  $c$ , where the expenditure distribution becomes too symmetric to account for the skewness of the log-normal. For high levels of homophily  $\rho$ , the parameter space for log-normality

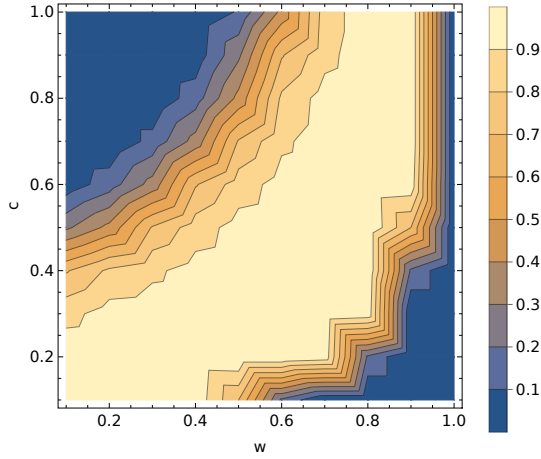


Figure 7.6: Fraction of expenditure distributions for which a standard Kolmogorov-Smirnov test cannot reject the assumption of log-normality at a 5% significance level for  $\rho = 0.5$  with 100 iterations.

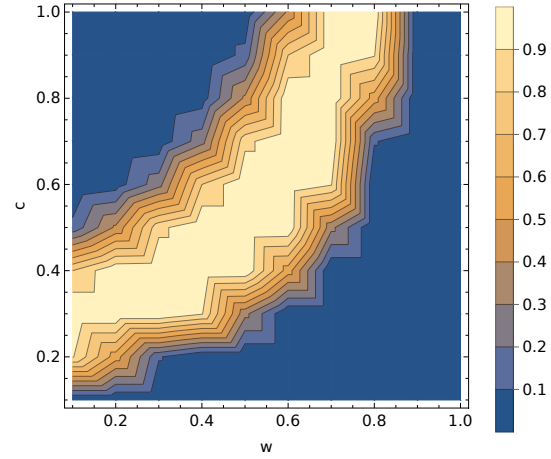


Figure 7.7: Fraction of expenditure distributions for which a standard Kolmogorov-Smirnov test cannot reject the assumption of log-normality at a 5% significance level for  $\rho = 4$  with 100 iterations.

becomes much narrower due to the emergent segregation that weakens the transmission of status consumption. In light of these findings, it seems entirely unsurprising that many of the empirical studies we reviewed in section 2 find significant deviations from log-normality in expenditure distributions, which one can understand as arising from variation in consumption baskets' salience of expenditures as well as differences in social segregation. Both the heterogeneity in the salience of consumption goods (Solnick and Hemenway 2005; Heffetz 2011, for a large scale study) and the variation in (e.g., spatial) segregation (Tóth et al. 2021) are well documented in the literature, in agreement with the varying results on the log-normality of expenditures across studies. Perhaps more surprisingly, we find that our stylised consumption function embedded within realistic social networks can replicate the relevant empirical phenomena in the literature on empirical expenditure patterns, testifying our modelling approach's validity.

#### 7.4.2 Inequality and Aggregate Savings

So far, the model assumed exponentially distributed incomes with a constant theoretical Gini coefficient of  $G = 0.5$ , irrespective of its particular parametrisation. To examine the impact of *changing* inequality on consumption and savings, we therefore now assume log-normality of individual incomes to generate variation in inequality through changing the dispersion parameter  $\sigma$  of the log-normal income distribution by which the model is initialised. We examine different degrees of homophily and let social networks adjust endogenously to input inequality, thereby highlighting two different findings: Firstly, there exist strong non-linearities in the relation of inequality and savings whenever perception networks mediate actual inequality. Secondly, the time scale of adjustment might be

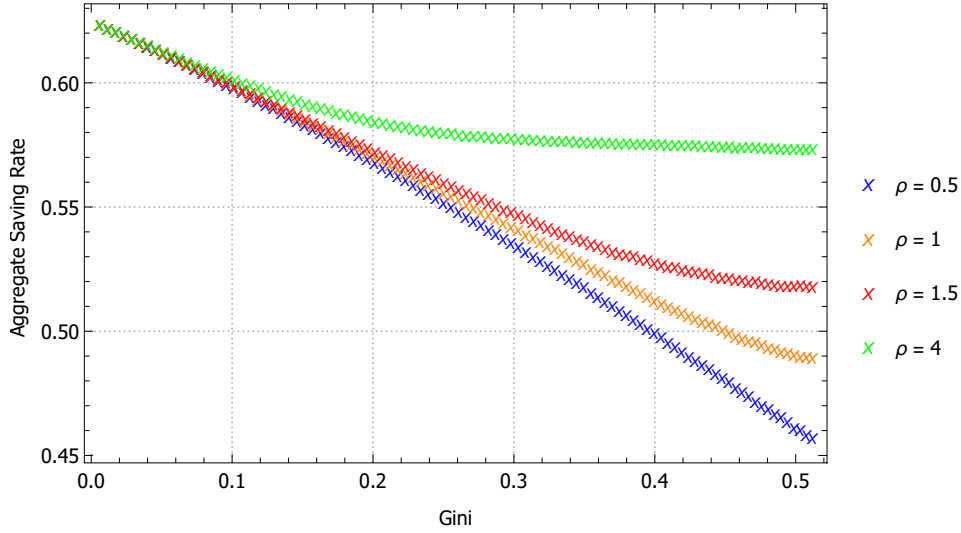


Figure 7.8: Aggregate saving rate (aggregate savings over aggregate income) as a function of inequality, measured by the Gini coefficient with  $w = c = 0.5$  and  $b = 0.75$ , averaged over 100 simulation runs.

important. We only provide comparative statics here and report results after all adjustments regarding perceptions have taken place but hint at possible mechanisms in the time domain as well.

Figure 7.8 depicts aggregate saving rates as a function of income inequality, measured by the Gini coefficient, and for varying levels of income homophily. As expected within a model of purely upward-looking consumption, savings unanimously decrease in inequality, as lower-income classes try to emulate the higher expenditures of the rich caused by their disproportionate gains in income. We are thus able to replicate the main finding of the canonical model of ‘expenditure cascades’ by Frank et al. (2014) within plausible and endogenously evolving social networks. However, we also find that effect sizes are both state-dependent for the very same behavioural parameters and decrease with homophily because for the given graph-generating process, two counteracting effects mediate the effect of actual inequality on perceived inequality (Schulz et al. 2022). Firstly, holding the network-topology constant, increases in aggregate inequality also increase inequality *within* perception groups and thus decrease savings. Secondly, the network topology endogenously mitigates an increase in inequality, too: This increasing inequality leads to higher segregation, which means higher inequality *between* groups but decreasing inequality within every given group. A priori, it is unclear what the relative strengths of both effects are, which we explore by simulation.

For (implausibly) low degrees of inequality, the four trajectories for  $\rho = 0.5, 1, 1.5$  and 4 almost appear to coincide, as the between-effect is negligible here, while the within-effect is fully passed through to savings. From a Gini coefficient  $G = 0.1$  onwards and now within the empirically relevant range (Bofinger and Scheuermeyer 2019), the trajectories



begin to deviate, with still strongly negative effects for the low-homophily regimes but vanishing savings elasticities within the highly homophilic regime. This is apparent from the decaying schedules for  $\rho = 0.5, 1$  and  $1.5$  in contrast to the plateau for  $\rho = 4$ . Methodologically, this finding thus points to a paradox of aggregation: With completely identical behavioural rules at the micro-level, one can generate vastly different aggregate savings elasticities at the macro level. Reminiscent to Kirman's (1992) foundational demonstration, the properties even of identical micro agents do not translate to equivalent properties of a 'representative agent' whatsoever, as the nonlinear macro behaviour in Figure 7.8 in contrast to the identically linear micro consumption rule shows. Albeit very different in detail, this result testifies to the relevance of the complex interaction of meso-level rules for aggregate consumption that Foster (2021) finds to be of crucial relevance for aggregate consumption patterns.

This methodological finding is not only a theoretical curiosum but also bears material implications: Most importantly, it might in part explain why the micro literature appears to confirm the relevance of the conspicuous consumption channel, while the macro literature remains at best ambiguous on this issue. It might very well be the case that all individuals engage in conspicuous consumption without this being observable using only macro-level aggregates, whenever there is an endogenous selection of reference groups. Indeed, the scarce extant empirical literature suggests both that inequality increases (geographical) segregation (Tóth et al. 2021) and that (geographical) segregation decreases conspicuous consumption (Bertrand and Morse 2016). Policies aimed at decreasing segregation might thus carry unintended consequences in simultaneously increasing conspicuous and potentially wasteful consumption.

Finally, our results are also a cautionary tale against the pitfalls of pooling macro data for analysis without acknowledging the potential time- and cross-sectional variation of inequality effects on segregation and perceptions. Indeed, it appears likely that the time-scale of the homophily mechanism outlined in section 3 crucially depends on the mobility of actors that enables homophilic choice or induced homophily in the first place (Thomas 2019, for a recent empirical study). With low mobility, the transitory adjustment to the new stationary state we depict in our comparative statics exercise in Figure 7.8 might take sufficiently long to be of macroeconomic relevance. In contrast, adjustment might be very fast for high degrees of mobility. The significant short-term elasticities, e.g. in Petach and Tavani (2021) can therefore be reconciled with the insignificant long-term effects Wildauer and Stockhammer (2018) find by considering heterogeneous mobility in contrast to ad-hoc assumptions regarding behavioural parameters.

Furthermore, the degree of homophily itself appears to be varying across and within countries (McPherson et al. 2001; Cepić and Tonković 2020). Consequently, pooled estimates might obscure unobserved cross-country heterogeneity both in adjustment speed and strength, which might help to explain the non-monotonicity Bofinger and Scheuer-

meyer (2019) find. Therefore, our model also calls for better data to capture these effects, either by controlling for, e.g., residential or occupational segregation or by directly including measures of perceived incomes. Unfortunately, these attempts are most of the time hindered by data that is much too coarse-grained for proper estimation, with the geographical location only being available at the state or county level (Bertrand and Morse 2016) and perceived incomes being reconstructed from data on a rather crude ordinal scale (Choi 2019) or based merely on ordinary workplaces (De Giorgi et al. 2020). Merging the relationship network, including friends and family, besides coworkers with the theoretically relevant socioeconomic data is rarely achieved in practice but would be an important step to further bridge social network and consumption theory (De Paula 2017).

## 7.5 Discussion

Our goal was to provide a tractable evolutionary alternative to orthodox theories of consumption based on Euler equations and intertemporal optimisation. In particular, we built on the distinction between ‘needs’ and ‘wants’ (Witt 2001) and integrated social norms through a perception network (Witt 2010). Our results suggest that this simple consumption rule based merely on current income and explicit observables can replicate the same set of major stylised facts as the orthodox formulation based on intertemporal consumption smoothing. In this sense, the two consumption theories appear to be observationally equivalent, and thus, even microdata on expenditure distributions might be unable to discriminate between the two. Theory, even in the form of carefully calibrated macro models, might be *underdetermined* in this regard (Quine 1975). Intertemporal motives might matter for consumption, but the log-normality of expenditures is perhaps not as decisive argument for consumption smoothing as it is often perceived to be (Battistin et al. 2009). Methodologically, intersubjective rather than intertemporal consumption rules might also be appealing since they build on much less restrictive assumptions. Intertemporal optimisation essentially implies both quantifiable income risk rather than fundamental uncertainty in the form of a well-defined probability distribution over income shocks and demands that this distribution is learnable or transparent for consumers (Menz 2010). By contrast, social consumption rests on cross-sectional observables made explicit by the perception network and therefore requires agents to be mere observers rather than econometricians which strikes us as a more plausible portrayal of human behaviour (Nelson and Consoli 2010). At the same time, our micro-level mechanisms are consistent with the permanent income hypothesis, too. By acting on what they see in others they perceive to be similar to themselves, individuals behave as if they were acting on expectations about their own lives. Hence, using social networks relying on interpersonal consumption comparisons might be attractive for future research as a proxy

for intertemporal considerations and its own sake because conspicuous consumption itself is an empirically well-grounded phenomenon.

Our second major result addresses the aggregation of individual consumption within the perception network. We find that, paradoxically, aggregate consumption might react far less elastic to changes in income than each individual consumption rule considered separately would indicate, even when all individuals unanimously follow exactly the same rule. This apparent contradiction follows from the endogenous selection of consumption reference groups. Apart from increasing aggregate consumption, inequality also increases segregation and therefore mitigates its initial consumption-enhancing effect. Naturally, this new channel sheds light on recent puzzles concerning the apparent divide between the micro and macro strands of the empirical literature on this matter. In particular, we explicate the potential micro origins of the institutional channel that Behringer and van Treeck (2018) and Ascione and Schnetzer (2022) find to be of crucial relevance to explain the empirical cross-country heterogeneity and time variation: In addition to the ease of access to credit or consumerist social norms, also the social network of everyday interactions might be an important determinant for the relative importance of expenditure cascades. The growing literature on ‘growth models’ within comparative political economy might thus also benefit from considering institutional differences in social segregation (Behringer and van Treeck 2019).

Moreover, our results point to a relevant trade-off for policy-makers: According to our model results, policies aimed at reducing occupational or residential segregation might also increase conspicuous consumption that might be environmentally wasteful (Howarth 1996) and contributes to the destabilising build-up of private debt (Van Treeck 2014). To offset these effects, our framework suggests complementing such policies of integration by appropriate taxation or behavioural nudges that are disincentivising status consumption (Antinyan et al. 2020), or, in the language of our model, decrease  $c$ . Note that this channel should be independent of our specific choice of consumption rule and qualitatively carries over to models with different utility functions, in particular, ratio instead of additive utility<sup>14</sup>, as long as there exists an upward-looking social consumption component as well as perception networks that are homophilic and adjust endogenously.

The choice in favour of a unanimous consumption rule based on comparisons is a deliberate one for two reasons: Firstly, our parsimonious model’s internal validation was unproblematic as we could monitor the whole parameter space of all exogenous variables and did not detect any inexplicable irregularities. Secondly, we keep individual behaviour intentionally simple to show that the *interaction* of agents is sufficient to generate empirical heterogeneity in micro observables and a large variety of macro-level effects. By applying a minimal model of consumption, we identify a (possible) minimal set of necessary assumptions (Grüne-Yanoff 2009). Hence, our model constitutes a how-possibly

---

<sup>14</sup>Cf. the discussion in Alpizar et al. (2005).

explanation of the empirically observed consumption patterns. However, the explanans can be true in the real world and the cause for the observed empirical fact: Our proposed mechanism fulfils the minimum conditions for a good epistemically possible how-possibly explanation formulated by Grüne-Yanoff and Verreault-Julien (2021). That our simulation results are in “qualitative agreement with empirical macrostructures” (Fagiolo et al. 2019, p. 771), namely successful replication of the stylised empirical facts, affirms the external validity of the model. Put differently, we develop a specific parallel reality (Sugden 2009) that features generating mechanisms for empirical findings in our reality and hence our results present a candidate explanation for the stylised empirical facts (Epstein 1999). Consequently, there may be different, more adequate, parallel realities featuring either these or even better mechanisms, despite to the best of our knowledge, there are no existing models that fulfil these characteristics. Furthermore, our proposed framework allows for arbitrary changes in effect sizes and micro patterns even without social consumption. To achieve this, one simply needs to alter individual consumption parameters ex-ante, e.g., by letting poorer individuals idiosyncratically consume a larger portion of their income and, as a corollary, increase the effect of inequality on aggregate consumption. Finally, both the principle of homophilic choice and social consumption effects are well established in the extant empirical literature, further testifying to the resemblance (Mäki 2009) between our model and reality. Therefore, our model reconciles technical verification with both input and descriptive output validation which Gräbner (2018) considers a rare feature of models. We are confident that the how-possibly explanation given by our model is preferable to ad-hoc choices and much more general or could at least be straightforwardly extended in this direction.

Nonetheless, the minimal nature of our model points to several limitations and extensions. Most notably, in its current bare-bones form, the model cannot study any issues of liquidity constraints that are frequently found to be of crucial relevance, especially for low-income households (Dynan et al. 2004), or spill-overs of private consumption to other sectors that would, at a minimum, require the model to incorporate general equilibrium effects and the model to be stock-flow consistent (Rengs and Scholz-Wäckerle 2019). Any supply-side factors or advertising activity to socially mediate ‘wants’ following Witt (2010) are also absent within our model that might nevertheless be of importance to explain the *changes* within consumption baskets and habits over time that Foster (2021) identifies. In particular the stock-flow consistent framework by Rengs and Scholz-Wäckerle (2019) or similar approaches could serve here as a starting point to incorporate our considerations within a full-fledged macro model.

Generally, links are modelled abstractly and incorporate various types of personal or professional connection layers as proximity of incomes. Future work could disentangle this conglomerate and represent specific ties more explicitly to focus on a particular connection layer, e.g., connections between co-workers implying network communities that represent

job categories. Early contributions to development economics on analogous ‘demonstration effects’ have emphasised the role of international emulation of consumption patterns (Nurkse 1953), while our perspective is much more granular and focuses on individuals. Exploring the interplay of aggregate frames of reference between countries and emulation of individual consumption within them might prove to be instructive to investigate cross-country convergence since the two could reinforce each other.

Furthermore, since the model is static, time does not play any role within the model per construction. Simon’s (1962) classic cautions us that the dynamic adjustment of different complex systems in response to exogenous disturbances might also operate on vastly different time scales, dependent on its decomposability properties. In this sense, the persistence of the effect of inequality on savings is a function of the time difference between consumption adjustment and network reshuffling. A possible remedy for this could be to explicitly model both consumption decisions, constrained by the availability of credit, and homophilic choice, constrained by mobility, within a full-fledged macro model. One promising avenue for empirical determination of the relevant time scales could be exploiting the time-variation in urbanisation rates. Veblen (1899) conjectured already in 1899 that cities are a powerful tool for diversifying perceptions, since within them, “the mobility of the population is greatest” (Veblen 1899, p. 66). The yet relatively scarce empirical evidence for Vietnam and Central and Eastern Europe suggests that perceived income diversity is indeed increasing with urbanisation (Mahajan et al. 2014; Binelli and Loveless 2016). Apart from theoretical models, we hope to invite empirical studies, especially at the intersection of urban or spatial economics, inequality research and consumer theory.

## References

- Agarwal, S., Mikhed, V., and Scholnick, B. (2020). Peers’ income and financial distress: Evidence from lottery winners and neighboring bankruptcies. *Review of Financial Studies*, 33(1):433–472.
- Akerlof, G. A. (1997). Social distance and social decisions. *Econometrica*, 65(5):1005–1027.
- Alpizar, F., Carlsson, F., and Johansson-Stenman, O. (2005). How much do we care about absolute versus relative income and consumption? *Journal of Economic Behavior & Organization*, 56(3):405–421.
- Alvarez-Cuadrado, F. and Van Long, N. (2011). The relative income hypothesis. *Journal of Economic Dynamics and Control*, 35(9):1489–1501.
- Annen, K. (2003). Social capital, inclusive networks, and economic performance. *Journal of Economic Behavior & Organization*, 50(4):449–463.

- Antinyan, A., Horváth, G., and Jia, M. (2020). Curbing the consumption of positional goods: Behavioral interventions versus taxation. *Journal of Economic Behavior & Organization*, 179:1–21.
- Anufriev, M., Chernulich, A., and Tuinstra, J. (2018). A laboratory experiment on the heuristic switching model. *Journal of Economic Dynamics and Control*, 91:21–42.
- Anufriev, M. and Hommes, C. (2012). Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics*, 4(4):35–64.
- Ascione, F. and Schnetzer, M. (2022). Out of balance? Revisiting the nexus of income inequality, household debt and current account imbalances after the Great Recession. *Cambridge Journal of Economics*, 46(1):133–165.
- Battistin, E., Blundell, R., and Lewbel, A. (2009). Why is consumption more log normal than income? Gibrat’s law revisited. *Journal of Political Economy*, 117(6):1140–1154.
- Behringer, J. and Van Treeck, T. (2018). Income distribution and the current account. *Journal of International Economics*, 114:238–254.
- Behringer, J. and van Treeck, T. (2018). Varieties of capitalism and growth regimes: The role of income distribution. *IMK Working Paper*, 193.
- Behringer, J. and van Treeck, T. (2019). Income distribution and growth models: a sectoral balances approach. *Politics & Society*, 47(3):303–332.
- Bertrand, M. and Morse, A. (2016). Trickle-down consumption. *Review of Economics and Statistics*, 98(5):863–879.
- Bianchi, M. and Sanfilippo, E. (2015). Social interdependencies in consumption: An early economic debate on social distinction, emulation, and fashion. *History of Political Economy*, 47(4):547–575.
- Binder, M. and Pesaran, M. H. (2001). Life-cycle consumption under social interactions. *Journal of Economic Dynamics and control*, 25(1-2):35–83.
- Binelli, C. and Loveless, M. (2016). The urban–rural divide: Perceptions of income and social inequality in Central and Eastern Europe. *Economics of Transition*, 24(2):211–231.
- Bofinger, P. and Scheuermeyer, P. (2019). Income distribution and aggregate saving: A non-monotonic relationship. *Review of Income and Wealth*, 65(4):872–907.
- Bordo, M. D. and Meissner, C. M. (2012). Does inequality lead to a financial crisis? *Journal of International Money and Finance*, 31(8):2147–2161.
- Boucher, V. (2015). Structural homophily. *International Economic Review*, 56(1):235–264.
- Bricker, J., Krimmel, J., and Ramcharan, R. (2020). Signaling status: The impact of relative income on household consumption and financial decisions. *Management Science*, 67(4):1–17.

- Brzozowski, M., Gervais, M., Klein, P., and Suzuki, M. (2010). Consumption, income, and wealth inequality in Canada. *Review of Economic Dynamics*, 13(1):52–75.
- Carr, M. D. and Jayadev, A. (2015). Relative income and indebtedness: Evidence from panel data. *Review of Income and Wealth*, 61(4):759–772.
- Carroll, C. D., Overland, J., and Weil, D. N. (1997). Comparison utility in a growth model. *Journal of Economic Growth*, 2(4):339–367.
- Cepić, D. and Tonković, Ž. (2020). How social ties transcend class boundaries? Network variability as tool for exploring occupational homophily. *Social Networks*, 62:33–42.
- Chakrabarti, A. S., Chatterjee, A., Nandi, T., Ghosh, A., and Chakraborti, A. (2018). Quantifying invariant features of within-group inequality in consumption across groups. *Journal of Economic Interaction and Coordination*, 13(3):469–490.
- Chandra, A. (2000). Labor-market dropouts and the racial wage gap: 1940–1990. *American Economic Review*, 90(2):333–338.
- Chen, W.-H., Myles, J., and Picot, G. (2012). Why have poorer neighbourhoods stagnated economically while the richer have flourished? Neighbourhood income inequality in canadian cities. *Urban Studies*, 49(4):877–896.
- Choi, G. (2019). Revisiting the redistribution hypothesis with perceived inequality and redistributive preferences. *European Journal of Political Economy*, 58:220–244.
- Christen, M. and Morgan, R. M. (2005). Keeping up with the Joneses: Analyzing the effect of income inequality on consumer borrowing. *Quantitative Marketing and Economics*, 3(2):145–173.
- Clementi, F. and Gianmoena, L. (2017). Modeling the joint distribution of income and consumption in Italy: A copula-based approach with  $\kappa$ -generalized margins. In Galletti, M., Palestini, A., and Russo, A., editors, *Introduction to Agent-Based Economics*, pages 191–228. Elsevier.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94:S95–S120.
- Cuaresma, J. C., Kubala, J., and Petrikova, K. (2018). Does income inequality affect aggregate consumption? Revisiting the evidence. *Empirical Economics*, 55(2):905–912.
- De Giorgi, G., Frederiksen, A., and Pistaferri, L. (2020). Consumption network effects. *Review of Economic Studies*, 87(1):130–163.
- De Paula, A. (2017). Econometrics of network models. In *Advances in Economics and Econometrics: Theory and Applications: Eleventh World Congress*, volume 1, pages 268–323. Cambridge University Press.
- Dos Santos, P. L. (2017). The principle of social scaling. *Complexity*, 2017.
- Drăgulescu, A. and Yakovenko, V. M. (2001). Evidence for the exponential distribution of income in the USA. *European Physical Journal B*, 20(4):585–589.

- Drechsel-Grau, M. and Schmid, K. D. (2014). Consumption–savings decisions under upward-looking comparisons. *Journal of Economic Behavior & Organization*, 106:254–268.
- Duesenberry, J. S. (1949). *Income, saving, and the theory of consumer behavior*. Harvard University Press.
- Dynan, K. E., Skinner, J., and Zeldes, S. P. (2004). Do the rich save more? *Journal of Political Economy*, 112(2):397–444.
- Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, 4(5):41–60.
- Fagiolo, G., Alessi, L., Barigozzi, M., and Capasso, M. (2010). On the distributional properties of household consumption expenditures: The case of Italy. *Empirical Economics*, 38(3):717–741.
- Fagiolo, G., Guerini, M., Lamperti, F., Moneta, A., and Roventini, A. (2019). Validation of agent-based models in economics and finance. In Beisbart, C. and Saam, N. J., editors, *Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives*, pages 763–787. Springer International Publishing, Cham.
- Falk, A. and Knell, M. (2004). Choosing the Joneses: Endogenous goals and reference standards. *Scandinavian Journal of Economics*, 106(3):417–435.
- Fixman, M., Berenstein, A., Brea, J., Minnoni, M., Travizano, M., and Sarraute, C. (2016). A bayesian approach to income inference in a communication network. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 579–582. IEEE.
- Foster, J. (2021). The US consumption function: a new perspective. *Journal of Evolutionary Economics*, 31(3):773–798.
- Frank, R. H. (1985). *Choosing the right pond: Human behavior and the quest for status*. Oxford University Press.
- Frank, R. H., Levine, A. S., and Dijk, O. (2014). Expenditure cascades. *Review of Behavioral Economics*, 1(1–2):55–73.
- Franke, R. and Westerhoff, F. (2017). Taking stock: A rigorous modelling of animal spirits in macroeconomics. *Journal of Economic Surveys*, 31(5):1152–1182.
- Friedman, M. (1957). *Theory of the consumption function*. Princeton University Press.
- Galaskiewicz, J. (2007). Has a network theory of organizational behaviour lived up to its promises? *Management and Organization Review*, 3(1):1–18.
- Ghosh, A., Gangopadhyay, K., and Basu, B. (2011). Consumer expenditure distribution in India, 1983–2007: Evidence of a long Pareto tail. *Physica A: Statistical Mechanics and its Applications*, 390(1):83–97.



- Gräbner, C. (2018). How to relate models to reality? An epistemological framework for the validation and verification of computational models. *Journal of Artificial Societies and Social Simulation*, 21(3).
- Grüne-Yanoff, T. (2009). Learning from minimal economic models. *Erkenntnis*, 70(1):81–99.
- Grüne-Yanoff, T. and Verreault-Julien, P. (2021). How-possibly explanations in economics: anything goes? *Journal of Economic Methodology*, 28(1):114–123.
- Gu, X., Dong, B., and Huang, B. (2015). Inequality, saving and global imbalances: A new theory with evidence from OECD and Asian countries. *The World Economy*, 38(1):110–135.
- Gu, X. and Huang, B. (2014). Does inequality lead to a financial crisis? Revisited. *Review of Development Economics*, 18(3):502–516.
- Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, 86(6):971–987.
- Hamilton, M. J., Milne, B. T., Walker, R. S., Burger, O., and Brown, J. H. (2007). The complex structure of hunter–gatherer social networks. *Proceedings of the Royal Society B: Biological Sciences*, 274(1622):2195–2203.
- Harting, P. and Radi, D. (2020). Residential segregation: The role of inequality and housing subsidies. *Journal of Economic Behavior & Organization*, 178:801–819.
- Heffetz, O. (2011). A test of conspicuous consumption: Visibility and income elasticities. *Review of Economics and Statistics*, 93(4):1101–1117.
- Hendry, D. F. and Mizon, G. E. (2014). Unpredictability in economic analysis, econometric modeling and forecasting. *Journal of Econometrics*, 182(1):186–195.
- Hoffman, S. D. and Duncan, G. J. (1988). Multinomial and conditional logit discrete-choice models in demography. *Demography*, 25(3):415–427.
- Hohnisch, M., Pittnauer, S., and Chakrabarty, M. (2002). Empirical regularities in distributions of individual consumption expenditure. *International Journal of Modern Physics C*, 13(04):541–549.
- Howarth, R. B. (1996). Status effects and environmental externalities. *Ecological Economics*, 16(1):25–34.
- Hu, X. and Liang, C.-Y. (2022). Does income redistribution prevent residential segregation? *Journal of Economic Behavior & Organization*, 193:519–542.
- Huckfeldt, R. R. (1983). Social contexts, social networks, and urban neighborhoods: Environmental constraints on friendship choice. *American Journal of Sociology*, 89(3):651–669.
- Jappelli, T. and Pistaferri, L. (2010). Does consumption inequality track income inequality in Italy? *Review of Economic Dynamics*, 13(1):133–153.

- Jappelli, T. and Pistaferri, L. (2014). Fiscal policy and MPC heterogeneity. *American Economic Journal: Macroeconomics*, 6(4):107–36.
- Jenkins, D. (2016). Conspicuous consumption in the United States and China. *Journal of Economic Behavior & Organization*, 127:115–132.
- Karlsson, N., Gärling, T., Dellgran, P., and Klingander, B. (2005). Social comparison and consumer behavior: When feeling richer or poorer than others is more important than being so. *Journal of Applied Social Psychology*, 35(6):1206–1222.
- Keynes, J. M. (1972). *Essays in Persuasion: The Collected Writings*, vol. 9 London. Macmillan.
- Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives*, 6(2):117–136.
- Klein, M. (2015). Inequality and household debt: a panel cointegration analysis. *Empirica*, 42(2):391–412.
- Koliofotis, V. (2022). Sexual selection of conspicuous consumption. *Journal of Evolutionary Economics*, 32:531 – 552.
- Koo, J. and Song, Y. (2016). The relationship between income inequality and aggregate saving: an empirical analysis using cross-country panel data. *Applied Economics*, 48(10):892–901.
- Krieger, N. (1992). Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *American Journal of Public Health*, 82(5):703–710.
- Krueger, D. and Perri, F. (2006). Does income inequality lead to consumption inequality? Evidence and theory. *Review of Economic Studies*, 73(1):163–193.
- Kuznets, S. (1942). *Uses of National Income in Peace and War*. National Bureau of Economic Research, New York.
- Lazarsfeld, P. F. and Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. In Berger, M., Abel, T., and Charles, H., editors, *Freedom and control in modern society*, pages 18–66. D. Van Nostrand Company.
- Leigh, A. and Posso, A. (2009). Top incomes and national savings. *Review of Income and Wealth*, 55(1):57–74.
- Leo, Y., Fleury, E., Alvarez-Hamelin, J. I., Sarraute, C., and Karsai, M. (2016). Socioeconomic correlations and stratification in social-communication networks. *Journal of The Royal Society Interface*, 13(20160598).
- Li, K.-H. and Li, C. T. (2019). Linear combination of independent exponential random variables. *Methodology and Computing in Applied Probability*, 21(1):253–277.
- Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, 15(3):215–233.

- Mac Carron, P., Kaski, K., and Dunbar, R. (2016). Calling Dunbar’s numbers. *Social Networks*, 47:151 – 155.
- Mahajan, S., Dinh, V. T., Rab, H., Trieu, V. Q., Nguyen, L. V., Kwakwa, V., and Verghis, M. (2014). Taking stock: an update on Vietnam’s recent economic developments. Technical report, Washington, D.C.: World Bank Group.
- Mäki, U. (2009). MISSing the world. Models as isolations and credible surrogate systems. *Erkenntnis*, 70:29–43.
- Malacarne, T. (2017). Rich friends, poor friends: Inter-socioeconomic status friendships in secondary school. *Socius*, 3:1 – 13.
- Manski, C. F. and McFadden, D. (1981). *Structural analysis of discrete data with econometric applications*. MIT press Cambridge, MA.
- Mayer, A. and Puller, S. L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics*, 92(1-2):329–347.
- Mayerhoffer, D. M. and Schulz, J. (2022a). Marginalisation and misperception: Perceiving gender and racial wage gaps in ego networks. In Benito, R. M., Cherifi, C., Cherifi, H., Moro, E., Rocha, L. M., and Sales-Pardo, M., editors, *Complex Networks & Their Applications X*, pages 768–779, Cham. Springer International Publishing.
- Mayerhoffer, D. M. and Schulz, J. (2022b). Perception and privilege. *Applied Network Science*, 7(1):1–25.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.
- Meghir, C. (2004). A retrospective on Friedman’s theory of permanent income. *Economic Journal*, 114(496):F293–F306.
- Menz, J.-O. (2010). Uncertainty, social norms and consumption theory: Post and New Keynesian approaches. *European Journal of Economics and Economic Policies: Intervention*, 7(1):125–146.
- Nelson, R. R. and Consoli, D. (2010). An evolutionary theory of household consumption behavior. *Journal of Evolutionary Economics*, 20(5):665–687.
- Newman, M. E. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409.
- Nurkse, R. (1953). *Problems of capital formation in underdeveloped countries*. Oxford University Press.
- Palley, T. I. (2010). The relative permanent income theory of consumption: A synthetic Keynes–Duesenberry–Friedman model. *Review of Political Economy*, 22(1):41–56.
- Perugini, C., Hölscher, J., and Collie, S. (2016). Inequality, credit and financial crises. *Cambridge Journal of Economics*, 40(1):227–257.
- Pesaran, M. H. (2015). *Time series and panel data econometrics*. Oxford University Press.

- Pesaran, M. H. and Weale, M. (2006). Survey expectations. In Granger, C. W. J., Elliott, G., and Timmermann, A., editors, *Handbook of economic forecasting*, volume 1, pages 715–776. Elsevier.
- Petach, L. A. and Tavani, D. (2021). Consumption externalities and growth: Theory and evidence for the United States. *Journal of Economic Behavior & Organization*, 183:976 – 997.
- Quine, W. V. (1975). On empirically equivalent systems of the world. *Erkenntnis*, 9(3):313–328.
- Quintana-Domeque, C. and Wohlfart, J. (2016). “Relative concerns for consumption at the top”: An intertemporal analysis for the UK. *Journal of Economic Behavior & Organization*, 129:172–194.
- Reardon, S. F. and Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, 116(4):1092–1153.
- Rengs, B. and Scholz-Wäckerle, M. (2019). Consumption & class in evolutionary macroeconomics. *Journal of Evolutionary Economics*, 29(1):229–263.
- Schmid, K. D. and Drechsel-Grau, M. (2013). Habits and envy: What drives the consumption behavior of us households? Evidence from PSID, 1999-2009. *IMK Working Paper Series*, 123.
- Schmidt-Hebbel, K. and Serven, L. (2000). Does income inequality raise aggregate saving? *Journal of Development Economics*, 61(2):417–446.
- Schulz, J., Mayerhoffer, D. M., and Gebhard, A. (2022). A network-based explanation of inequality perceptions. *Social Networks*, 70:306 – 324.
- Sen, A. (1984). The living standard. *Oxford Economic Papers*, 36:74–90.
- Silva, A. C. and Yakovenko, V. M. (2004). Temporal evolution of the “thermal” and “superthermal” income classes in the USA during 1983–2001. *Europhysics Letters (EPL)*, 69(2):304.
- Simon, H. (1962). The architecture of complexity. *Proceedings of American Philosophical Association*, 106(6).
- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Republished (1991), Everyman’s Library Classics Series, London: Home University Library.
- Smith, J. A., McPherson, M., and Smith-Lovin, L. (2014). Social distance in the United States. *American Sociological Review*, 79(3):432–456.
- Solnick, S. J. and Hemenway, D. (2005). Are positional concerns stronger in some domains than in others? *American Economic Review*, 95(2):147–151.
- Sugden, R. (2009). Credible worlds, capacities and mechanisms. *Erkenntnis*, 70(1):3–27.

- Tao, Y., Wu, X., Zhou, T., Yan, W., Huang, Y., Yu, H., Mondal, B., and Yakovenko, V. M. (2019). Exponential structure of income inequality: Evidence from 67 countries. *Journal of Economic Interaction and Coordination*, 14(2):345–376.
- Thomas, R. J. (2019). Sources of friendship and structurally induced homophily across the life course. *Sociological Perspectives*, 62(6):822–843.
- Toda, A. A. (2017). A note on the size distribution of consumption: More double Pareto than lognormal. *Macroeconomic Dynamics*, 21(6):1508–1518.
- Tóth, G., Wachs, J., Di Clemente, R., Jakobi, Á., Ságvári, B., Kertész, J., and Lengyel, B. (2021). Inequality is rising where social network segregation interacts with urban topology. *Nature Communications*, 12(1):1–9.
- Uzzi, B. and Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology*, 111(2):447–504.
- Van Treeck, T. (2014). Did inequality cause the US financial crisis? *Journal of Economic Surveys*, 28(3):421–448.
- Veblen, T. (1899). *The Theory of the Leisure Class*. Republished (2001), New York: The Modern Library.
- Virtanen, M., Kivimäki, M., Kouvonen, A., Elovainio, M., Linna, A., Oksanen, T., and Vahtera, J. (2007). Average household income, crime, and smoking behaviour in a local area: The Finnish 10-town study. *Social Science & Medicine*, 64(9):1904–1913.
- Watson, T. (2009). Inequality and the measurement of residential segregation by income in American neighborhoods. *Review of Income and Wealth*, 55(3):820–844.
- Weeden, K. A. and Cornwell, B. (2020). The small-world network of college classes: implications for epidemic spread on a university campus. *Sociological Science*, 7:222–241.
- Wheeler, C. H. and La Jeunesse, E. A. (2008). Trends in neighborhood income inequality in the US: 1980–2000. *Journal of Regional Science*, 48(5):879–891.
- Wilcox, K. and Stephen, A. T. (2013). Are close friends the enemy? Online social networks, self-esteem, and self-control. *Journal of Consumer Research*, 40(1):90–103.
- Wildauer, R. and Stockhammer, E. (2018). Expenditure cascades, low interest rates, credit deregulation or property booms? Determinants of household debt in OECD countries. *Review of Behavioral Economics*, 5(2):85–121.
- Witt, U. (2001). Learning to consume—a theory of wants and the growth of demand. *Journal of evolutionary economics*, 11(1):23–36.
- Witt, U. (2010). Symbolic consumption and the social construction of product characteristics. *Structural Change and Economic Dynamics*, 21(1):17–25.
- Xu, Y., Santi, P., and Ratti, C. (2021). Beyond distance decay: Discover homophily in spatially embedded social networks. *Annals of the American Association of Geographers*, 112(2):505–521.

- Yanev, G. P. (2020). Exponential and hypoexponential distributions: Some characterizations. *Mathematics*, 8(12):2207.
- Zhou, W.-X., Sornette, D., Hill, R. A., and Dunbar, R. I. (2005). Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B: Biological Sciences*, 272(1561):439–444.

# Aggregate Consumption, Perception Networks and Functional Distribution

---

Revised version submitted with D. M. Mayerhoffer as “Individual perception, consumption emulation, and demand regimes” to the Review of Political Economy.

I want to thank Anna Gebhard, Leonhard Ipsen, Daniel Mayerhoffer, Malcolm Sawyer, Daniele Tavani, Jan Weber, as well as the participants of the fourth Behavioral Macroeconomics Workshop in Bamberg and the 26th Forum for Macroeconomics and Macroeconomic Policies in Berlin for many helpful comments and questions.

---

## Abstract

Empirical (personal) income distributions robustly follow a two-class structure, with an exponential bulk of labour income accruing to the lower 95 - 99 % population share and an upper Pareto tail of the 1 - 5% richest, where capital income is concentrated (Silva and Yakovenko 2004). The implications of this regularity for macroeconomic outcomes, especially aggregate consumption, are not yet fully understood, though. We introduce this two-class structure into a Post-Keynesian consumption model with workers and capitalists corresponding to the two classes within the personal distribution. Agents consume according to idiosyncratic and social motives like in Mayerhoffer and Schulz (2021); status consumption is microfounded within a perception network as discussed in Schulz et al. (2022) that replicates empirically observed inequality and social self-perceptions. Our findings indicate that the non-market interactions in perception networks are potentially highly relevant for aggregate outcomes and that they can shed some light on recent puzzles and open questions in the Post-Keynesian literature on growth regimes. In particular, we show how upward-looking consumption externalities can be made consistent with wage-led consumption and how network segregation can explain differences in the degree of ‘wage-ledness’ of aggregate consumption. Our findings carry two implications: First, the effects of shifts in the functional distribution of income might well depend on social norms and the *perceived* inequality within social networks. Second, empirical studies should take the regularities both in the income distribution as well as social network topologies into account, with aggregate representations being potentially misleading.

---

---

## Contents of this Paper

8.1	Introduction . . . . .	233
8.2	Model . . . . .	234
8.2.1	Consumption Function . . . . .	234
8.2.2	Perception Network . . . . .	236
8.2.3	Income Distribution . . . . .	237
8.2.4	Sequence of events . . . . .	241
8.3	Results . . . . .	241
8.3.1	Median and Mean Consumption Norm . . . . .	241
8.3.2	Micro: Distributional Stylised Facts . . . . .	242
8.3.3	Macro: Median Consumption Norm . . . . .	246
8.3.4	Macro: Mean Consumption Norm . . . . .	248
8.4	Discussion . . . . .	250
	References . . . . .	253

---



## 8.1 Introduction

The foundational contribution by Bhaduri and Marglin (1990) has spawned a vast literature on growth regimes and the possibility of wage-led growth. This is unsurprising, since wage-led growth implies a synergy rather than a trade-off between the two goals of equity and prosperity, rendering the predictions of the neo-Kaleckian model immensely relevant for economic policy. While the literature has largely focused on elaborations of the investment function or open-economy models (Lavoie 2022), the consumption function has not attracted similar attention and is still largely based on the original premise of heterogeneous consumption propensities that are higher for workers than for capitalists. Redistribution from profits to wages thus unanimously increases aggregate consumption within those models. This theoretical assumption of heterogeneous consumption propensities is so well established empirically (Dynan et al. 2004) that contradictory results are dubbed ‘anomalous’ or ‘perverse’ (Barbosa-Filho and Taylor 2006, table 1; Stockhammer and Stehrer 2011, table 2; Stockhammer 2017).<sup>1</sup> Our model is an attempt to go beyond class-based heterogeneous consumption propensities and show that accounting for interaction and emulative motives between heterogeneous agents enriches the picture. We find that both wage- and profit-led consumption regimes can emerge within our synthesis model and that decreasing income-based segregation both increases aggregate consumption levels and makes the growth regime more wage-led. The main contribution of this paper is to explicate the relevance of non-market interactions by showing that they exhibit significant influence on aggregate variables with several examples.

We address two apparent empirical paradoxes within heterodox consumption theory: First, we explore the puzzle that consumption functions featuring upward-looking status comparisons should always imply profit-led consumption (Schulz et al. 2022), in contrast to the overwhelming majority of empirical studies that show consumption to be wage-led (cf. e.g. Onaran and Obst 2016). This is due to the fact that increases in the profit share theoretically and empirically should generally also imply increased inequality in the personal distribution of income (Atkinson 2009). This, in turn, increases aggregate consumption in models of upward-looking status comparisons. We show that this result in Schulz et al. (2022) depends on the *type* of inequality increase and only holds unanimously, whenever the increase results from within-class rather than between-class increases in inequality. The problem can be remedied by introducing the standard Kaleckian assumption of heterogeneous consumption propensities between workers and capitalists. Consequently, the wage-ledness of consumption depends on the relative population shares of capitalists and workers and the degree of income homophily within the

---

<sup>1</sup>Note that the foundational paper by Barbosa-Filho and Taylor (2006) does not discuss the ‘anomaly’ of their estimated consumption function. Yet, table 1 clearly shows that consumption responds negatively to an increase in the wage share in contrast to the Kaleckian argument. For a related discussion, cf. Kapeller and Schütz (2015).

perception networks, though, not only on the different consumption propensities. Second, these factors might help to interpret the recent aforementioned “perverse” results regarding the implied profit-ledness of consumption in several empirical studies (Barbosa-Filho and Taylor 2006; Stockhammer and Stehrer 2011; Stockhammer 2017), in violation of the prediction of the Bhaduri-Marglin framework (Bhaduri and Marglin 1990). By contrast, our framework can indeed feature both profit- and wage-led consumption.

Theoretically, we also demonstrate that heterogeneous consumption propensities between workers and capitalists are neither necessary nor sufficient to generate wage-led consumption: Consumption can be wage-led for homogeneous marginal consumption propensities, and it can be profit-led even in the presence of such heterogeneity. We show this result to be dependent primarily on the network topology of perceptions that represents real world day-to-day interactions. Network segregation generally makes consumption more profit-led, since expenditure cascades within the working class become weaker and thus consumption out of profits grows relatively more important.

The remainder of this paper is organised as follows. Section 8.2 introduces the model and discusses the validation strategy for the various underlying assumptions it uses. Section 8.3 discusses our main results both regarding the distributional regularities the model generates on a micro-level as well as the distribution-led regimes for different parametrisations and degrees of network segregation. The final section 8.4 concludes and discusses policy implications and limitations of the model.

## 8.2 Model

Our model setup consists of three distinct building blocks: i) A consumption function with heterogeneous, class-based parameters, ii) the perception network that determines reference consumption for all agents and iii) the income distribution by which we initialise both the perception network and the consumption function. We introduce and discuss all three building blocks in turn.

### 8.2.1 Consumption Function

All agents, indexed  $i \in \{1, \dots, N\}$ , consume according to a consumption function with identical functional form. The individual consumption expenditure  $C_i$  by agent  $i$  is given by

$$C_i = w_i \cdot Y_i + (1 - w_i) \cdot c_i (C(j|i) - w_i \cdot Y_i), \quad (8.1)$$

where  $Y_i$  denotes current disposable income and reference consumption  $C(j|i)$  is the highest consumption level by agent  $j$  that agent  $i$  observes in their perception network. Schulz

and Mayerhoffer (2022) show that this consumption rule emerges from utility maximisation for Cobb-Douglas preferences that penalises deviations from reference consumption  $C(j|i)$ . The parameter  $w_i$  is then the elasticity parameter for consumption expenditures and  $(1 - w_i)$  for savings to capture intertemporal motives, while  $c \geq 0$  determines the intensity of (linear) disutility from falling short of the reference consumption level.

While the above consumption rule can be derived from a marginalist approach with a clear interpretation of all its parameters, it is consistent with Post-Keynesian consumer choice, too, that motivated the model in the first place. In this interpretation, the consumption rule would be an expression of “procedural rationality” (Simon 1976) which Lavoie (1994) considers to be a central tenet of Post-Keynesian consumer theory. The given consumption rule thus explicates a rule of thumb that allows consumers to decide quickly and efficiently. The first term would correspond to idiosyncratic consumption with  $w_i$  as the marginal propensity to consume (MPC) out of current disposable income. By contrast, the second term includes social motives and captures status consumption. Finally, the social term follows what Lavoie (1994) calls “non-independence”. Consumers have a “desire to impress the Joneses, which makes each family strive to keep up at least an appearance of being as well off as those that they mix with” (Robinson 1956, p. 251). The upward-lookingness of consumption is also empirically well established (Bertrand and Morse 2016; Bricker et al. 2020; Heffetz 2011; Jinkins 2016; Quintana-Domeque and Wohlfart 2016; Drechsel-Grau and Schmid 2014; Alpizar et al. 2005). Like in the above consumption function, consumption “thus reflects the lifestyle of the households that constitute its social reference group” (Eichner 1986, p. 160). The functional form of the consumption function thus follows a common Post-Keynesian vision of consumer choice. We also follow the Post-Keynesian vision regarding the *parameter choices* for the consumption function in eq. (8.1). In particular, we consider two classes, workers and capitalists (indexed *work* and *cap*). Parameters are homogeneous within but not between classes. Regarding the MPC out of current disposable income, we employ the conventional Kaleckian assumption that MPCs out of profits are much lower than those out of wages (Kalecki 1971; Lavoie 1992), which implies

$$0 \leq w_{cap} \leq w_{work} \leq 1. \quad (8.2)$$

This assumption is empirically verified, cf. e.g. Dynan et al. (2004). Note that the above relationship allows for the possibility of workers’ savings like in Harris (1974) and Mott and Slatiery (1994) as a generalisation of the original argument by Kalecki (1971) who abstracted from any savings by workers. Regarding the social component, we set

$$c_{work} > 0 \text{ and} \quad (8.3)$$

$$c_{cap} = 0. \quad (8.4)$$

The assumption that only workers engage in conspicuous consumption is entirely conventional (perhaps to sustain tracatability) (Kapeller and Schütz 2015; Setterfield and Kim 2017; Detzer 2018) but also justifiable empirically: Stanley and Danko (1996) show that the super rich do not engage in conspicuous consumption and, if anything, practice conspicuous *frugality*<sup>2</sup> or explicitly *inconspicuous consumption*, as is also shown for the domain of fashion by Davis (1992) and Brooks (2010). Sherman (2018) and Friedman and Reeves (2020) confirm this finding and show that the modes of distinction for the super rich have in today’s Western cultures shifted towards less conspicuous pursuits. Charles et al. (2009) demonstrate empirically that belonging to a relatively poor group (in our case, the working class) induces conspicuous consumption, presumedly to fend off the perception of “looking poor”.

### 8.2.2 Perception Network

Like in Schulz and Mayerhoffer (2022), the perception of workers is constrained by a perception network following Schulz et al. (2022) and Mayerhoffer and Schulz (2022).<sup>3</sup> They show that applying this random geometric graph type of model generates perception graphs that can efficiently reproduce the documented stylised facts about overall inequality perceptions as well as perceptions of intergroup inequities. They also show that the emergent network topologies are consistent with real world social networks, i.e., they feature small-worldiness.

Given these externally validated properties, this type of graph is a reasonable way to replicate the real-world interactions of workers. The major ingredient of the model is income homophily, the empirical tendency of individuals with similar income to form ties to each other (McPherson et al. 2001). Following Mac Carron et al. (2016), each agents draws five other agents to link to. Draws are weighted by income distance, i.e., Agent  $j$ ’s weight in agent  $i$ ’s draw is denoted by  $\Omega_{ij}$  and determined as follows:

$$\Omega_{ij} = \frac{1}{\exp[\rho |Y_j - Y_i|]} \quad (8.5)$$

---

<sup>2</sup>For example, they show that 70% of American millionaires in their samples purchase cars with below average market value.

<sup>3</sup>Capitalists are not included within the network since they do not engage in status consumption. This ensures that we can isolate the effects of the *aggregate* measure of capitalists’ consumption and do not need to worry about confounding effects by individual capitalists’ consumption that are used as reference consumption. This assumption is equivalent to assuming that the two classes do not interact directly with each other.

The relative weights in the draws are a function of the homophily strength  $\rho \in \mathbb{R}_0^+$  and the respective income levels  $Y_j$  and  $Y_i$ . Schulz and Mayerhoffer (2022) and Schulz et al. (2022) show that the above linkage formation process can be microfounded within a random utility framework and that the weights in eq. (8.5) can be translated into probabilities  $\omega$  as follows:

$$\omega_{ij} = \frac{\exp[-\rho \cdot |Y_j - Y_i|]}{\sum_{k \in W \setminus i} \exp[-\rho \cdot |Y_k - Y_i|]}, \quad (8.6)$$

with  $W \setminus i$  as the set of all workers except  $i$ . For any given income distribution of discrete income levels  $Y$ , the linkage probabilities are thus fully determined by the homophily strength. It is easy to see that for  $\rho = 0$ , the graph reduces to an Erdős-Rényi random graph (Erdős and Rényi 1960) but for increasing  $\rho > 0$ , the network grows more and more segregated as choice becomes more homophilic in income. We are just building on the empirically well-established correlation<sup>4</sup> in reduced form but are silent on the specific *cause* of emergent segregation. In particular, we do not assume that income homophily is necessarily the product of specific choices of individuals and allow for *induced* homophily by shared workplaces or educational stratification (McPherson and Smith-Lovin 1987; Chetty et al. 2022a). While both types can be represented by the same homophily parameter  $\rho$  and thus have the same effect on simulation outcomes, they might carry drastically different policy implications, as we summarise in section 8.4.

Finally, we need to close the model and specify, how the two classes interact with each other. To this end, we assume that the reference consumption level of the *richest worker* indexed  $k$ ,  $C(j|k)$  is given by a measure of central tendency of the consumption of capitalists, i.e., either their mean  $\bar{C}_{cap}$  or median consumption  $C_{cap}^{median}$ . Here, we follow the intuition that workers plausibly do not observe the consumption patterns of particular capitalists but rather a general picture that is mediated by social networks (Wang et al. 2017) or television (Ger and Belk 1999). As it turns out, the choice between the mean and median consumption as the reference level has drastic aggregate consequences.<sup>5</sup>

### 8.2.3 Income Distribution

As the final building block, we need to determine the income distribution by which the perception network and consumption formation is initialised. Here, we build on the

---

<sup>4</sup>Cf. Schulz and Mayerhoffer (2022), Schulz et al. (2022) and Mayerhoffer and Schulz (2022) for rather comprehensive reviews of the empirical literature

<sup>5</sup>Apart from these contentual reasons, this assumption is also necessary to guarantee the internal validity of the model: Pseudorandom draws from a power law with small sample sizes are notoriously problematic and vary dramatically between samples (Talaga and Nowak 2020, Appendix C). Since we want to study out the effect of the profit share on aggregate consumption rather than the noise of the random number generator, we work with aggregate measures instead that do not vary between draws.

well-established finding of a two-class income distribution that is characterised by an exponentially distributed bulk for the vast majority and a power law upper tail of the 1 - 5 % super rich (Silva and Yakovenko 2004; Kumar 2017; Shaikh 2017; Shaikh and Jacobo 2020; Kumar 2021).<sup>6</sup> We follow Silva and Yakovenko (2004) and Shaikh and Jacobo (2020) who show that the exponential bulk of the distribution consists primarily of labour income (i.e., wages), while the Pareto tail is populated by capitalists earning capital income. Consequently, we identify the exponential part with the labor share of national income and the Pareto tail with the profit share. With some fairly general technical assumptions, this then allows to fully characterise the income distribution by the population share of capitalists  $p$  and the aggregate mark-up over unit labor costs  $m$  as well as the total population size  $N$ . We thus provide a mapping from the continuous income distribution to categorical classes.

Like in the standard Kaleckian framework, the model world is characterized by oligopolistic competition. In particular, we assume that the aggregate mark-up  $m$  on unit labour costs is exogenous, leading to the following equation for the economy-wide average price level  $P$  as

$$P = (1 + m) \frac{W}{Y} \quad (8.7)$$

with  $W$  as the the total wage bill and  $Y$  as total output. We are thus abstracting from raw material costs for simplicity (Hein 2014, p. 247). The profit share  $h$  of total income can thus be derived as

$$h = \frac{\Pi}{PY} = \frac{PY - W}{PY} = \frac{(1 + m)W - W}{(1 + m)W} = \frac{m}{1 + m}, \quad (8.8)$$

where  $\Pi$  are aggregate profits (Hein 2014, chps. 6 and 7).<sup>7</sup> The profit share  $h$  is therefore uniquely determined by the average mark-up  $m$ . We assume that profits only accrue to the  $p$ -th upper percentile of the income distribution that marks the onset of the Pareto

---

<sup>6</sup>The findings by Silva and Yakovenko (2004) were heavily criticised by Schneider (2015) and Schneider and Scharfenaker (2020) for relying too heavily on graphical methods and thus failing to account for mixtures that include other components, in particular, a log-normal component. As a theoretical rationale, they propose labor market segmentation and dualisation following the seminal work by Reich et al. (1973). While being beyond the scope of this paper, it would in principle be possible to include such a mixture into the model here to study the effects of dualisation on aggregate consumption in a model extension.

<sup>7</sup>Since  $m$  is directly proportional to the profit share  $h$ , we could restate the whole sequence of derivations following this part in terms of  $h$ . We choose to state all variables in terms of  $m$ , though, to be compatible with the canonical model and to hopefully ease integration into a full-fledged macro model, where  $m$  is endogenous. Also, it is helpful to consider, what different profit shares imply in terms of market power: A profit share of 1/3, following the famous rule of thumb (Johnson 1954; Izyumov and Vahaly 2015), implies an aggregate markup of  $m = 1/2$ , while a profit share of 1/2 already implies that  $m = 1$ .

distribution. The lower part of the income distribution consisting of wages follows an exponential distribution with rate parameter  $\lambda$ , such that the mean of the exponential regime  $\bar{Y}_{work} = 1/\lambda$ . Note that with this assumption, interpersonal inequality as measured by the Gini coefficient always increases with the profit share  $h$  and the mark up  $m$ , replicating the finding in Atkinson (2009) derived in a non-parametric manner for a specific distributional regime.<sup>8</sup> This implies that we can express total income as

$$(1 - p) \cdot N \cdot \bar{Y}_{work} + \frac{m}{1 + m} N \bar{Y} = Y = N \cdot \bar{Y}. \quad (8.9)$$

Since we are only interested in shares, we can assume without loss of generality that the mean total income  $\bar{Y} = 1$ , such that  $Y = N$ , which implies

$$(1 - p) \cdot \bar{Y}_{work} + \frac{m}{1 + m} = 1, \quad (8.10)$$

and thus  $\bar{Y}_{work}$  can be expressed as

$$\bar{Y}_{work} = \frac{1}{(1 + m)(1 - p)}. \quad (8.11)$$

The mean income within the Pareto tail of capitalists is given by

$$\bar{Y}_{cap} = \frac{m}{1 + m} \cdot \frac{Y}{pN} = \frac{m}{1 + m} \cdot \frac{N}{pN} = \frac{1}{p} \frac{m}{1 + m}, \quad (8.12)$$

since there are  $pN$  incomes within the Pareto tail and  $Y = N$ .

If status consumption is  $c_{cap} = 0$  within the Pareto tail following assumption eq. (8.4), total consumption within the tail is thus given as

$$C_{cap} = w_{cap} \left( N \cdot \frac{m}{1 + m} \right). \quad (8.13)$$

We can also derive the implied tail exponent  $\alpha$  from the mean tail income  $Y_{cap}^{min}$  by setting the minimum income parameter of the  $Y_{cap}^{min}$  to the (expected) maximum of the expo-

---

<sup>8</sup>Silva and Yakovenko (2004) show that the Gini coefficient  $G$  for a two-class distribution of this kind can be approximated for a low population share within the power law regime as  $G = (1 + f)/2$ , with  $f \geq 0$  denoting the share of income within the power law tail with respect to total income. In our model,  $f$  is defined as the profit share  $m/(m + 1)$  which implies that  $G = 1 - 1/(2 \cdot (1 + m))$  with the expected behaviour. For  $m = 0$ ,  $G = 1/2$  which is simply the Gini coefficient of the exponential wage distribution, while for  $m \rightarrow \infty$ ,  $G \rightarrow 1$ , since then the income share of vanishingly small population share of capitalists approaches unity. Finally,  $G$  is unambiguously increasing in  $m$ , i.e., functional inequality translates into interpersonal inequality.

nential, i.e.,  $Y_{cap}^{min}$  marks the (expected) threshold income between the two distributional regimes. Order statistics of an exponential involve the number of draws from this distribution. Rényi (1953) shows that the expected value of the maximum of an exponential distribution with rate parameter  $\lambda$  after  $n$  draws is given by  $H_n/\lambda$ , with  $H_n$  being the  $n$ -th Harmonic number. We can thus determine the threshold income as

$$Y_{cap}^{min}(p, m, N) = H_{(1-p) \cdot N} \cdot \frac{1}{(1-p)(1+m)}. \quad (8.14)$$

The threshold income increases with  $N$  (as there is then a higher number of draws and the exponential density is unbounded from above) by the Harmonic number  $H_{(1-p) \cdot N}$ . Since the Harmonic number is well approximated by  $\log[n] + \gamma$  (with  $\gamma$  as the Euler–Mascheroni constant and  $\gamma = 0.5772\dots$  (Dence and Dence 2009)), the effect of a change in population size is logarithmic and thus rather small, though. Also, the higher the markup  $m$ , the higher the share of income accruing to profits for fixed other parameters and thus, the earlier the “onset” of the power-law regime in terms of income. Finally, there are two counteracting effects of an increase in  $p$ . Increasing  $p$  first decreases  $Y_{cap}^{min}$  due to a decrease in the number of draws from the exponential, as can be seen from the Harmonic number as the first factor. Second, however, it also decreases the number of people within the exponential bulk and thus ceteris paribus increases the mean income of workers. Since the Harmonic number decreases rather slowly with  $p$  (i.e., logarithmically) for  $p \ll 1$  as the first factor, while entering as (much faster) hyperbolic growth in the second term, the whole threshold income tends to increase in  $p$  which we also verify for our simulations. This then allows us to derive the implied tail exponent by equating the mean tail income derived purely from economic considerations above with the mean income implied by a Pareto distribution, i.e.,

$$Y_{cap}^{min} \frac{\alpha}{\alpha - 1} \stackrel{!}{=} \frac{1}{p} \frac{m}{1 + m} \quad (8.15)$$

$$\Leftrightarrow \alpha = \frac{m(p - 1)}{pH_{(1-p)N} + mp - m} \quad (8.16)$$

Using the quantile definition of a Pareto distribution  $F_{Pareto}^{-1}(q) = Y_{min}^{Pareto}(1 - q)^{-1/\alpha}$ , we can thus finally derive the implied median (or any other quantile) of this Pareto tail as

$$F_{Pareto}^{-1}(q; p, m, N) = \frac{H_{(1-p)N}(1 - q)^{\frac{pH_{(1-p)N}}{(1-p)m} - 1}}{(m + 1)(1 - p)}, \quad (8.17)$$

with the median at  $q = 1/2$ .



### 8.2.4 Sequence of events

Below, we show the proposed sequence of events to determine the aggregate APC for one Monte Carlo run:

1. Initialise  $(1 - p)N \in \mathbb{N}$  workers.
2. Allocate income according to exponential distribution with desired rate parameter.
3. Generate network for workers based on income distribution for given homophily strength  $\rho$ .
4. Initialise consumption of workers by determining their idiosyncratic consumption levels and setting the reference consumption  $C(j|k)$  of the richest worker either to the mean consumption within the tail  $w_{cap} \cdot \bar{Y}_{cap}$  or some consumption quantile  $w_{cap} \cdot F_{cap}^{-1}(q; p, m, N)$ .
5. Sequentially update the consumption levels of all workers going down the income distribution by always taking as reference consumption the highest observed consumption level. Sum over workers' consumption to get  $C_{work}$ .
6. The aggregate APC is then given by

$$\frac{C_{work} + C_{cap}}{Y} = \frac{C_{work} + w_{cap} \left( N \frac{m}{1+m} \right)}{N}. \quad (8.18)$$

## 8.3 Results

For analysis, we set all parameters within the empirically calibrated range<sup>9</sup>, i.e.,  $w_{work} = 0.5$ ,  $c_{work} = 0.5$ ,  $p = 0.02$  and  $N = 1,000$  and vary  $\rho \in [1; 14]$  and  $m \in [0.5, 1]$ . We consider first the empirical results on the micro level for external validity to show that the emergent consumption expenditure distributions indeed replicate empirical stylised facts. The goal of the macro results is to show that both the level and wage-ledness of aggregate consumption depend crucially on parameters beyond the idiosyncratic MPCs of each class.

### 8.3.1 Median and Mean Consumption Norm

Our first result is completely analytical. We show in Figures 8.1 and 8.2 that while the mean income of capitalists unsurprisingly increases with the profit share, the median

---

<sup>9</sup>We also conduct sensitivity analyses covering the whole parameter space. Results are available upon request.

income tends to decrease in it for  $p < 0.05$ . Thus, while the capitalist class as a whole benefits from a higher share of profits, the *typical* capitalist is left worse off. This result is reminiscent of the basic premise of ergodicity economics that economic variables are typically non-ergodic (Peters 2019) and that growth of the average income is thus often radically different from average or “typical” income growth (Adamou and Peters 2016; Adamou et al. 2020).<sup>10</sup>

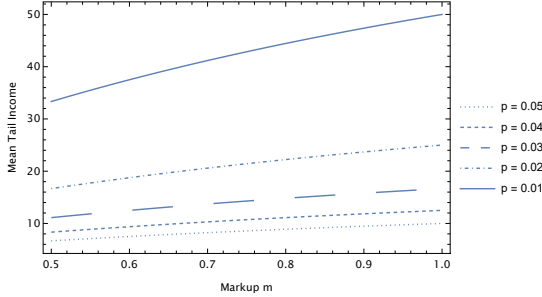


Figure 8.1: Mean income within the Pareto tail for varying  $p$  and  $m$ . Mean tail incomes unambiguously increase in markups and thus, profit shares.

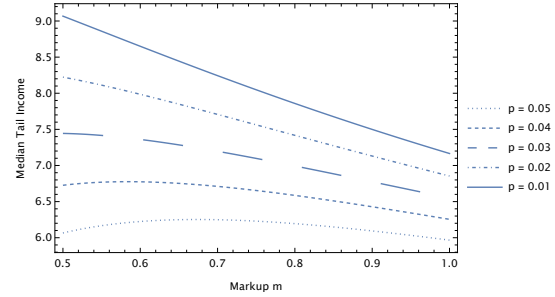


Figure 8.2: Median income within the Pareto tail for varying  $p$  and  $m$ . Median tail incomes tend to *decrease* in markups and thus, profit shares.

The reason for this behaviour is that an increase of the profit share has two counteracting effects on the median tail income: Firstly, by construction, it increases the total tail income and thus also median income. However, secondly, the minimum tail income  $Y_{min}^{cap}$  is defined as the (expected) threshold between the exponential and power law regime and thus depends negatively on the mean income in the exponential regime. As we show above, the second effect appears to dominate the first for (almost all) reasonable  $p$  and  $m$  and the median tail income declines in the profit share. As a direct corollary to this, tail inequality also increases with the profit share, as the mean income unambiguously increases with  $m$ , while the minimum tail income also unambiguously decreases with the profit share. Since capitalists’ consumption is directly proportional to their income (due to  $c_{cap} = 0$ ) the polar behaviour of the mean and median tail income has strong implications for the effect of the mean or median consumption norms the richest worker takes as reference consumption, as dicussed in section 8.2.2 and 8.2.3.

### 8.3.2 Micro: Distributional Stylised Facts

We validate our model by comparing our simulation outcomes to stylised empirical facts about consumption expenditure distributions in increasing level of detail and granularity

<sup>10</sup>Identical concepts were already introduced in Milanovic (2005) and Saez and Zucman (2019) as a “population-weighted average growth rate” and “people’s growth rate”, respectively.

following Fagiolo et al. (2019).<sup>11</sup> The first stylised fact concerns the decrease of individual APCs and is thus validated at level 1 of Fagiolo et al.'s (2019) taxonomy, i.e., it is qualitatively consistent with empirical macrostructures. First, and almost trivially by construction, we find that like in the real world, APCs tend to decline in income, i.e., the rich save more out of current income (Dynan et al. 2004; Clementi and Gianmoena 2017). This effect of expenditure cascades (Frank et al. 2014) is what we find consistently for

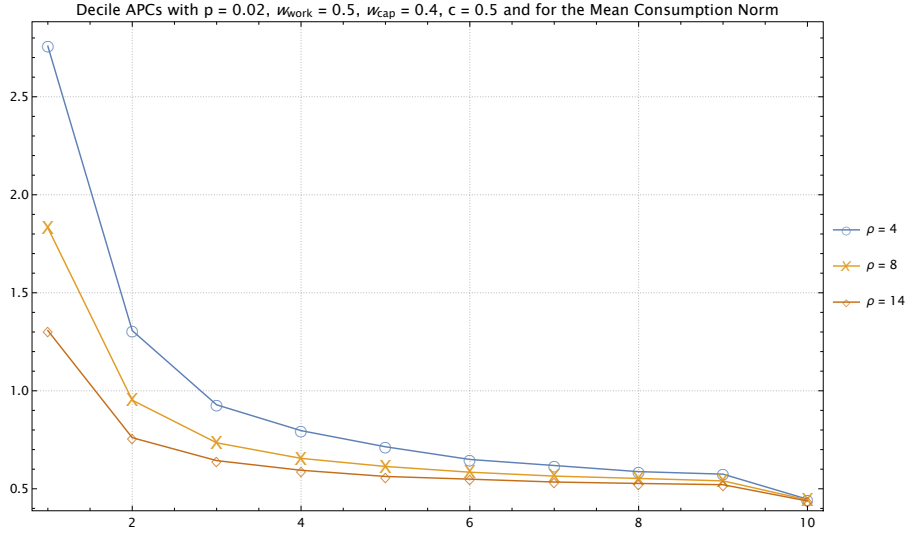


Figure 8.3: Decile level APCs for a single simulation run and  $w_{work} = 0.5, w_{cap} = 0.4, c = 0.5, p = 0.02$  and various  $\rho \in \{4; 8; 14\}$ . Lines are visual aids only. APCs consistently decline in income decile with APCs being higher the lower network segregation (measured by  $\rho$ ) is.

our model, as shown for a single parametrisation of the consumption function and different network topologies. Workers aim to catch up to higher levels of consumption they observe which implies that status consumption accumulates in its way down the income distribution with the poorest consequently exhibiting the highest share of status consumption. Thus, part of the decline of individual APCs is per assumption ( $w_{work} > w_{cap}$ ) but the decline within the working class follows from upward-looking status consumption and is endogenous to the model. This effect on the poorest varies with the network topology though, with the poorest decile being heavily affected for low levels of network segregation. This follows from the fact that low income homophily tends to increase the increment between observed (highest) consumption and own idiosyncratic consumption by construction. Thus, the more diverse the network contacts are on average, the stronger the expenditure cascade.

We also replicate the stylised fact that consumption expenditure distributions are empirically much more homogeneous than income distributions (Krueger and Perri 2006;

<sup>11</sup>All results are for the mean consumption norm but hold for the median consumption norm with other parametrisation as well.

Jappelli and Pistaferri 2010). The model is thus also validated at level 2 following Fagiolo et al. (2019), i.e., it is in *quantitative* agreement with empirical macrostructures. We show this below in Figure 8.4 only for a specific parametrisation ( $w_{work} = 0.5, c = 0.5, m = 0.8$  and  $\rho = 8$ ) but income distributions are consistently more heterogeneous than expenditure distributions for parametrisations we considered within the permitted range.

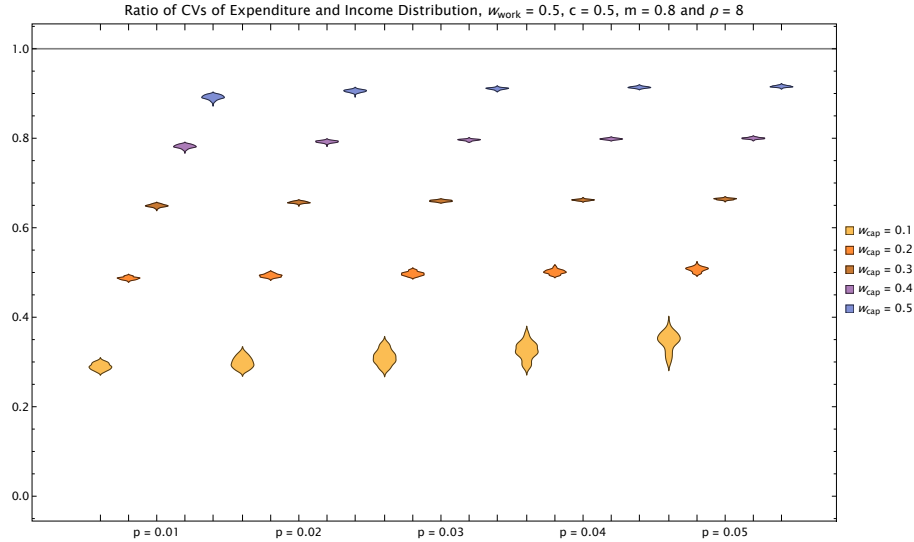


Figure 8.4: Violin plots of the ratios of the coefficients of variation (CV) of the consumption expenditure distributions to the coefficients of variation of the income distributions, i.e.  $CV_C/CV_Y$ . Values below unity (solid line) indicate that expenditure distributions are more homogeneous than income distributions, as is the case across all simulations and parametrisations.

We measure the variability of both distributions by their respective coefficients of variation. Figure 8.4 shows the ratios of coefficients of variation of  $CV_C$  over the coefficients of variation for the income distributions  $CV_Y$ . Ratios below unity thus indicate that the stylised fact is replicated. This is indeed the case, both for homogeneous and heterogeneous consumption propensities and for different population shares of capitalists and workers: All simulated ratios are below the grid line at unity. This is due to the fact that the poor catch up to the rich, as also indicated by their higher MPCs. The emergent expenditure cascades thus homogenise consumption. Heterogeneous marginal consumption propensities, i.e.,  $w_{work} > w_{cap}$  add to this homogenising tendency. Capitalists consume absolutely more than workers due to their higher income but this effect is mitigated, when they consume relatively less idiosyncratically. The lower their idiosyncratic MPC, the more homogeneous the expenditure distribution. This effect is visible in the Violin plots that indicate a clear increase in expenditure heterogeneity for increasing  $w_{cap}$ .

Finally, empirical studies find that the distribution of consumption expenditures is as a first-order approximation well fit by a log-normal distribution (Battistin et al. 2009; Brzozowski et al. 2010; Chakrabarti et al. 2018; Hohnisch et al. 2002; Fagiolo et al. 2010;

Ghosh et al. 2011; Toda 2017), while the distribution of current income is not (Silva and Yakovenko 2004; Kumar 2017; Shaikh and Jacobo 2020; Kumar 2021). Yet, there exist significant deviations from the log-normal benchmark, particularly for their upper tails (Chakrabarti et al. 2018; Fagiolo et al. 2010; Ghosh et al. 2011; Toda 2017) that often follow power laws. This is what we also consistently find for our expenditure distributions,

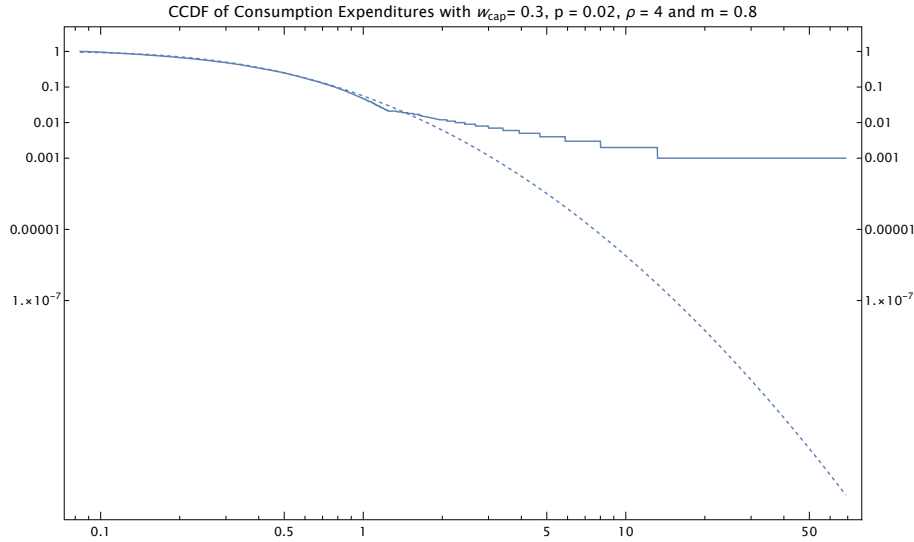


Figure 8.5: CCDF of consumption expenditures on double-logarithmic scale for  $p = 0.02$ ,  $w_{cap} = 0.3$ ,  $\rho = 4$ ,  $m = 0.8$  and  $c = 0.5$ . Superimposed dashed line is best log-normal fit for the data. Consumption expenditures are clearly more heavy-tailed than log-normal but are well approximated by log-normal density for lower strata of the distribution.

like e.g. Figure 8.5. The bulk of the distribution is well approximated by a log-normal that is superimposed on the CCDF but the distribution also clearly has a much heavier upper tail that is approximately linear on a double-log scale. We thus find a mixture of a log-normal for the vast majority with an upper Pareto tail which is the exact functional form that Ghosh et al. (2011) find for their dataset. Without explicitly targeting it, our model is thus able to make sense of the fact that consumption is more log-normal than income (Battistin et al. 2009) and that deviations from log-normality occur primarily in the upper tail (Chakrabarti et al. 2018; Fagiolo et al. 2010; Ghosh et al. 2011; Toda 2017). It is this heavy upper tail that models of intertemporal optimisation typically fail to replicate (Battistin et al. 2009). We thus also validate one property of our model at level 3 in Fagiolo et al. (2019), i.e., quantitative agreement with empirical microstructures. The generating mechanism for the log-normal part is described in Schulz and Mayerhoffer (2022). Our consumption rule coupled with exponentially distributed wages leads to an expenditure distribution that is hypoexponential and thus well approximated by a log-normal density (Yanev 2020). The consumption expenditures are simply a rescaled version by  $0 < w_{cap} < 1$  of the upper tail of the income distribution and thus follow a power law by construction.

Our model is thus validated both with respect to the input parameters, i.e., the income distribution and inequality perceptions that are in line with empirical data, and regarding the output distribution of consumption expenditures. The main model mechanisms taken in isolation are also already established in the literature, i.e., heterogeneous consumption propensities regarding consumption out of profit and wages (Dynan et al. 2004), the given perception network (Schulz et al. 2022) as well as upward-looking consumption behaviour (Heffetz 2011). Gräbner (2018) considers this a rare but desirable feature of models to be validated both in terms of input and output variables and mechanisms. Even though the model is rather stylised, it thus replicates an unexpectedly high number of stylised facts even at the highest level of validation that Fagiolo et al. (2019) considers. We thus consider the model sufficiently validated to at least replicate parts of empirical reality and turn to the macroeconomic outcomes it generates.

### 8.3.3 Macro: Median Consumption Norm

We first simulate the model for homogeneous marginal consumption propensities (i.e.,  $w_{work} = w_{cap} = 0.5$ ) and show the shares of aggregate consumption with respect to total income for 100 Monte Carlo runs in Figure 8.6. Consumption is wage-led even for homogeneous consumption propensities, with effect sizes quickly decreasing for high degrees of homophily. This is exactly the result we expect from our analytical derivation for the median and mean tail income: Increasing the profit share decreases both the median tail income and thus the reference consumption of the richest worker as well as the individual income and idiosyncratic consumption of all workers. Consumption expenditures of workers thus unambiguously decrease with an increasing profit share. By contrast, consumption expenditures of capitalists unambiguously increase with the profit share. This effect cannot compensate for the loss in workers' expenditures, though, since the APCs of capitalists are strictly below the ones of workers due to  $c_{cap} = 0$ . The effect appears to be stronger (and aggregate consumption levels are generally higher), when the increments between own and observed reference consumption are larger, i.e., when network segregation is lower. Workers tend to observe a more diverse set of other workers, whenever the homophily parameter is low which increases both the consumption increment and the emergent expenditure cascades. The standard Kaleckian assumption of heterogeneous consumption propensities is thus not necessary to generate wage-led consumption but for the effect of the profit share to be significant, perception networks need to be sufficiently heterogeneous.

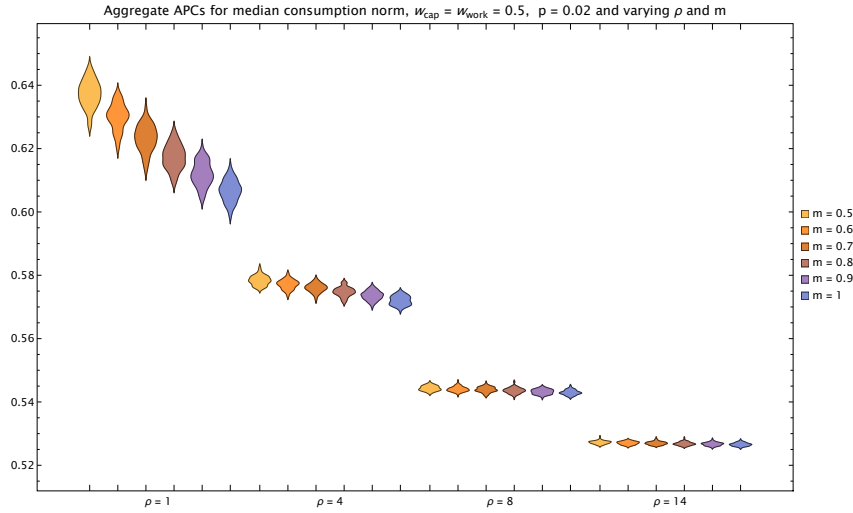


Figure 8.6: Violin plots of aggregate APCs of 100 Monte Carlo runs with  $p = 0.02$ , the consumption norm for the richest worker given by the median consumption of capitalists, identical consumption propensities  $w_{cap} = w_{work}$  and for varying  $m$  and  $\rho$ . Even for homogeneous consumption propensities, aggregate consumption is wage-led but the degree of “wage-ledness” gets much weaker for higher  $\rho$ .

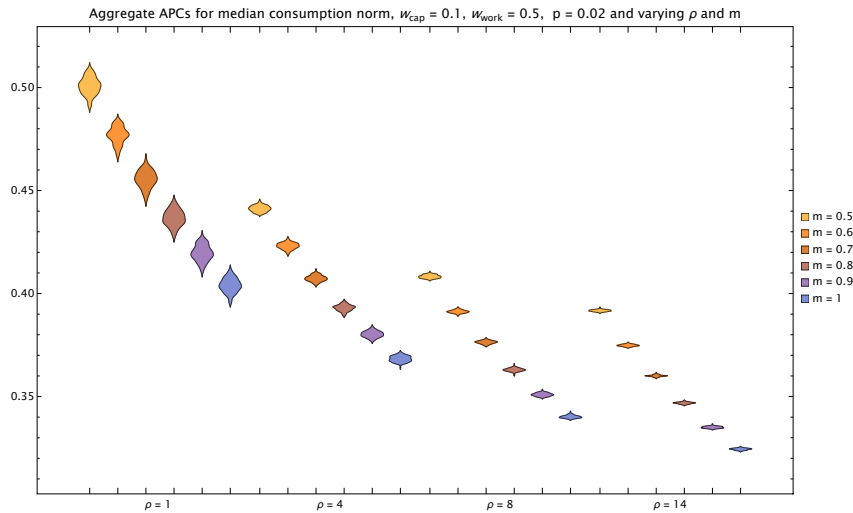


Figure 8.7: Violin plots of aggregate APCs of 100 Monte Carlo runs with  $p = 0.02$ , the consumption norm for the richest worker given by the median consumption of capitalists,  $w_{cap} \ll w_{work}$  and for varying  $m$  and  $\rho$ . For this large difference in consumption propensities, aggregate consumption is strongly wage-led.

Introducing the Kaleckian assumption of heterogeneous consumption propensities compounds the initial wage-ledness of consumption for these assumptions. Here, we let the idiosyncratic MPC of workers be 5 times large than the MPC of capitalists, i.e.,  $w_{work} = 0.5 \gg w_{cap} = 0.1$ . Figure 8.7 shows the results. Aggregate APCs decay almost linearly with markups for all given levels of  $\rho$ . Consumption is thus strongly wage-led

irrespective of homophily, with the consumption level again being inversely related to the level of network segregation. If reference consumption of the richest worker is given by median tail consumption for the given parameter restrictions, aggregate consumption is thus always and irrespective of network segregation wage-led. Yet, the level of marginal propensities to consume of both classes and the degree of income homophily determine the level of aggregate consumption, without affecting the growth regime regarding consumption.

### 8.3.4 Macro: Mean Consumption Norm

If reference consumption of the richest worker is determined by the average tail consumption, the growth regime responds to network segregation, though. Consider again the case of homogeneous consumption propensities with results summarised in Figure 8.8. Here, consumption undergoes a regime shift and is wage-led for low homophily parameters ( $\rho = 1$ ) but becomes profit-led for higher levels of network segregation ( $\rho \geq 4$ ). This

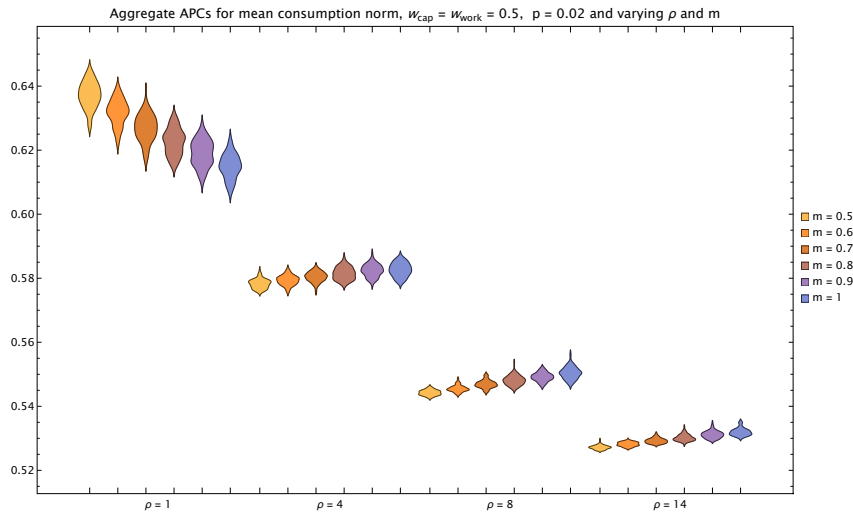


Figure 8.8: Violin plots of aggregate APCs of 100 Monte Carlo runs with  $p = 0.02$ , the consumption norm for the richest worker given by the mean consumption of capitalists and for varying  $m$  and  $\rho$ . For low levels of  $\rho$ , consumption is wage-led even for homogeneous consumption propensities due to relatively low network segregation.

result ultimately stems from differences in segregation at the top. As Schulz et al. (2022) show, the combination of exponential wage distribution and homophilic linkage leads to a large degree of segregation at the top of the wage distribution that increases disproportionately with  $\rho$ . For high levels of  $\rho$ , the number of workers that directly perceive the consumption of the richest worker is thus sufficiently high to let aggregate consumption be profit-led: The consumption of the richest worker  $k$  increases with the profit share due to the higher reference consumption. For high clustering at the top, this emulation effect for the workers that observe  $k$ 's consumption dominates the effect of decreased income



and consumption of all other workers from a lower wage share. For low  $\rho$ , clustering at the top is not high enough to compensate the decreased idiosyncratic consumption levels of all the workers other than  $k$ .<sup>12</sup> We thus show that the network topology can induce a regime shift in aggregate consumption and confirm the finding that heterogeneous MPCs are not a necessary condition for wage-led consumption.

As a direct corollary to this, we also show that not all inequalities are alike: Schulz and Mayerhoffer (2022) demonstrated that increases in interpersonal inequality within a class (i.e., increasing the dispersion parameter of a log-normal income distribution) unambiguously increases aggregate consumption for homogeneous parameters. By contrast, we show that increases of interpersonal inequality caused by changes of the functional distribution, i.e., an increase of the profit share, might very well decrease aggregate consumption. Summary measures of interpersonal (income) inequality such as the Gini coefficient might thus mask differences in the specific *type* of inequality considered. Empirical studies on the inequality-growth nexus might thus benefit from including measures that focus on top tail incomes or the functional distribution more generally rather than just the Gini coefficient (Berg et al. 2018; Flechtner and Gräbner 2019) that is rather insensitive to changes at the tails of the distribution (Atkinson 1970).

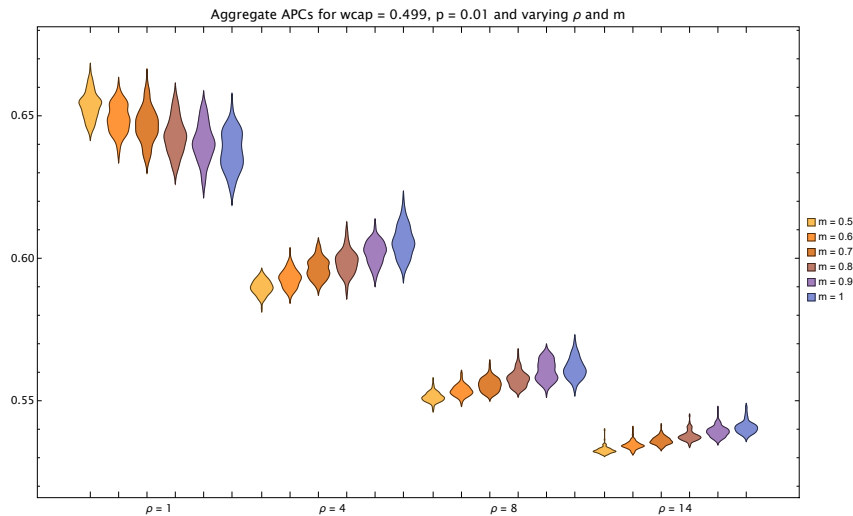


Figure 8.9: Violin plots of aggregate APCs of 100 Monte Carlo runs with  $p = 0.02$ , the consumption norm for the richest worker given by the mean consumption of capitalists and for varying  $m$  and  $\rho$ . For medium to high  $\rho$ , consumption is profit-led, even though consumption propensities of capitalists are (slightly) smaller.

We also show that heterogeneous consumption propensities are not sufficient to generate wage-led consumption, if the MPC differential is sufficiently small. In Figure 8.9, we show the simulation results for MPCs of capitalists being slightly lower than the ones

<sup>12</sup>We also verify this mechanism by monitoring the perception sets of workers for different  $\rho$ . Simulation results and the corresponding Julia codes are available upon request.

of workers, i.e.,  $w_{cap} = 0.499 < w_{work} = 0.5$  to illustrate this. As is clearly visible, consumption stays profit-led for medium to high degrees of homophily. Heterogeneous consumption propensities are thus neither necessary nor sufficient to generate wage-led consumption for our parameter restrictions. Much like in Kirman's (1992) classic, aggregate consumption behaviour is thus irreducible to micro consumption rules: Neither does wage-led consumption imply lower MPCs of capitalists nor do lower MPCs by capitalists imply wage-led consumption.

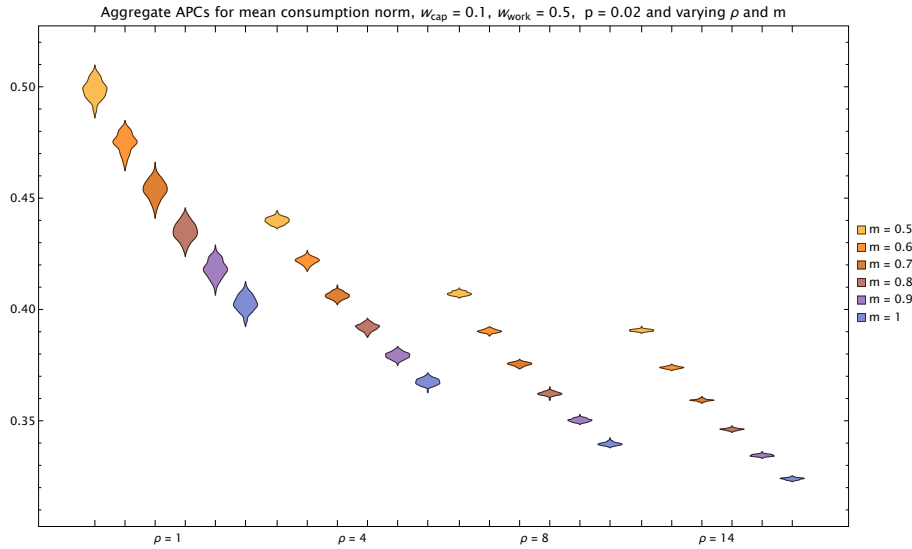


Figure 8.10: Violin plots of aggregate APCs of 100 Monte Carlo runs with  $p = 0.02$ , the consumption norm for the richest worker given by the mean consumption of capitalists, varying  $m$  and  $\rho$  and. For this large difference in  $w_{work} \gg w_{cap}$ , consumption  $C$  is strongly wage-led.

Finally, Figure 8.10 shows that aggregate consumption exhibits the expected pattern, whenever MPCs are highly heterogeneous, in this case  $w_{work} = 0.5 \gg w_{cap} = 0.1$ . If we empirically determine that the consumption behaviour differs strongly between classes, we are thus indeed justified to ignore perceptions and network effects. In this sense, our model recovers the original insight by Kalecki (1971) that consumption is wage-led, if workers do not save or, in the language of the well-known aphorism “spend what they get” (Gladys Foster 1990), while capitalists’ consumption propensity is much lower. If classes are similar or even equal in their idiosyncratic consumption behaviour, perceptions and interaction in networks become important, with even the growth regime shifting between network topologies.

## 8.4 Discussion

We introduce a parsimonious consumption model that synthesises the main insights of class-based heterogeneity following Kalecki (1971) and non-independence following e.g.

Robinson (1956) and Eichner (1986). Independent of the reference consumption norm for the richest worker, aggregate consumption (as a share of national income) is higher, the lower network segregation is (Schulz and Mayerhoffer 2022). This might be helpful to make sense of differences in the aggregate *level* of consumption which exhibit tremendous international heterogeneity<sup>13</sup> and are often ignored by studies that typically focus on the sign and effect size of the derivative of GDP with respect to the wage share. Focusing on the growth regime of aggregate consumption, we recover the original Kaleckian insight that consumption is wage-led, whenever the differential in consumption propensities is sufficiently large. The picture becomes much more nuanced, when the two classes converge in terms of consumption behaviour. In this case, both the network segregation and consumption norms become important. As a final theoretical result, we also show that the result in Schulz and Mayerhoffer (2022) that interpersonal inequality unanimously increases consumption has to be qualified: If the increase in interpersonal inequality results from within-class changes, aggregate consumption indeed increases but for increases that result from changes in the functional distribution, it is possible for aggregate consumption to *decrease* instead. Disentangling the two channels is thus also a promising direction for empirical studies.

If the reference consumption of the richest worker is given by the median consumption level of capitalists, consumption is unambiguously wage-led.<sup>14</sup> This is due to a curious analytical result that the *typical* capitalist might get poorer when the *average* capitalist benefits, reminiscent of the findings of Adamou and Peters (2016). The more interesting case is thus certainly the one where the reference consumption of the richest worker is the mean of capitalists' consumption. Here, we document the distinct possibility of profit-led consumption for empirically plausible parameter regions which might imply that the “anomalies” found by Barbosa-Filho and Taylor (2006); Stockhammer and Stehrer (2011) and Stockhammer (2017) might not be so anomalous after all and can be explained within the confines of our model. Our simulations also provide proof of concept that heterogeneity in consumption propensities are neither necessary nor sufficient to generate wage-led consumption.

Profit-led consumption tends to emerge for high degrees of network segregation and consumption behaviour that is rather homogeneous between classes. This points to an interesting policy conclusion: Decreasing network segregation might have three potentially beneficial effects. First, it tends to reduce bias in inequality perceptions and thus aids in getting more informed voting decisions (Schulz et al. 2022; Mayerhoffer and Schulz

---

<sup>13</sup>Values for the consumption share of GDP range from 15 % for Turkmenistan to 146 % for Somalia, with large differences even for OECD countries, e.g. Germany with only 49 % compared to 67 % for the US (World Bank 2022).

<sup>14</sup>It might also appear distribution-neutral for high levels of network segregation and homophily, since the effect of status consumption is greatly diminished here, leaving the impression that the profit share has no effect at all.

2022). Second, it tends to increase aggregate consumption and thus growth (Schulz and Mayerhoffer 2022), while it also appears to enhance economic mobility (Chetty et al. 2022b). And finally, it might provide the foundations for a strategy of wage-led growth, i.e., benefit policies that aim to simultaneously decrease inequality and foster economic growth. Here, the specific type of homophily becomes important that we both represent in reduced form by a single parameter: If homophily is primarily *induced* homophily, policy might arguably indeed decrease segregation by targeting the structural characteristics that cause it, e.g., by subsidising poorer neighbourhoods or by appropriate public housing programs. By contrast, for *choice* homophily one would need to influence the preferences and beliefs of people themselves which is perhaps the much harder task and would require educational campaigns.<sup>15</sup> Empirically, both types of homophily are likely interlinked (Kossinets and Watts 2009) and disentangling both types might provide an exciting new avenue for further research regarding *desegregation-led* growth.

This synergetic relationship between the three points above is evidently appealing for policy. Yet, it also highlights a potential shortcoming of the model: It completely abstracts from any intertemporal motives and, in particular, household debt. Expenditure cascades imply that individuals at the lower tail of the income distribution exhibit the highest APC and thus, bear the brunt of necessary debt to finance consumption expansions. This might have problematic distributional consequences (Frank et al. 2014) and contribute to financial fragility á la Minsky (Kapeller and Schütz 2015). Our model is blind for these causes of aggregate volatility and fragility.

While the focus on consumption appears justified due to increased relevance of consumption-led expansions (Kharroubi and Kohlscheen 2017), studying e.g. the effect of debt and liquidity constraints requires a full-fledged macro model that integrates other markets, in particular, financial markets (Kapeller and Schütz 2015). Also, the model completely abstracts from any kind of policy that might help to overcome instability, as is shown for monetary policy by Proaño et al. (2011). We aimed to make our partial consumption model compatible with such attempts to study spill-overs, e.g. by endogenising income and market power materialising itself in changes in national income and its distribution to ease integration within a macro framework. Finally, one recently documented phenomenon could also complicate the applicability of the rather strict assumptions of the model at hand: Homoploutia, the fact that the rich are increasingly both labor and capital income-rich (Berman and Milanović 2020; Iacono and Ranaldi 2021; Kumar 2021). The cut-off between labor and capital income that appeared reasonable following the find-

---

<sup>15</sup>Choice homophily is indeed a potential major factor for residential segregation according to race and refutes intuitions that people's preferences are typically only "mildly" racist or segregationist. The ALLBUS 2006 study for Germany included an item on the preferred degree of segregation with the vast majority of respondents preferring almost completely segregated neighbourhoods with only a few "foreigners" (GESIS – Leibniz-Institut für Sozialwissenschaften 2011). These individual determinants of segregation are likely to be overlooked when focusing only on structural characteristics.

ings of Silva and Yakovenko (2004) or Shaikh and Jacobo (2020) might become more and more fuzzy.<sup>16</sup> We will address the issues of liquidity constraints and *composition inequality* (Iacono and Ranaldi 2021) in further research.

The minimal nature of our model obviously points to various limitations which we discussed above. Yet, our findings indicate that the use of aggregate statistics might be severely misleading, when heterogeneity is large and behaviour is interdependent (Kirman 1992). We introduce both heterogeneity and interaction in form of the empirically validated income distribution and perception network and explicate several theoretical channels by which the functional distribution of income can influence aggregate consumption. Ultimately, the relevance of these channels can only be established empirically, though. We hope to prompt empirical investigation into the effect of network segregation on aggregate consumption. We thus underline the need to go beyond *actual* inequality and study the effects of *perceived* inequality, since perceptions and not necessarily the actual state of affairs shape behaviour. While precise measurement of network segregation, consumer norms and perceptions is a hard problem to solve empirically, considering those factors might enrich the literature and help to explain international variability in growth regimes and aggregate consumption shares. In particular, the more macroeconomically oriented literature on institutional differences in growth regimes might benefit from the theoretical mechanisms presented here to study the microeconomically heterogeneous effects of aggregate norms (Behringer and van Treeck 2018; Ascione and Schnetzer 2022). Chapter 11 of Shiller's (2019) *Narrative Economics* analyses the effects of norms and narratives on frugality and conspicuous consumption which illustrates how a focus on extra-economic norms and perceptions might aid in understanding (shifts in) growth regimes. In the terminology of Colander (1993), the two-class structure of income distributions and perception network thus might provide the *macrofoundations* of consumption behaviour that is *microfounded* with upward-looking consumption behaviour on the micro level that is itself endogenous to consumer norms. Our main contribution with this model is to show that both cannot be studied in isolation from each other.

## References

- Adamou, A., Berman, Y., and Peters, O. (2020). The two growth rates of the economy. *arXiv preprint*, 2009.10451.
- Adamou, A. and Peters, O. (2016). Dynamics of inequality. *Significance*, 13(3):32–35.

---

<sup>16</sup>Note however that this phenomenon of homoploutia might be partially explainable by the recent surge in CEO payments (Gabaix and Landier 2008) that are more 'akin to profits' (Kalecki 1971, p. 76). In particular, the consumption behaviour by CEOs might resemble the one of more traditional capitalists and CEO payments will typically be located within the Pareto tail. The sharp cut-off might thus continue to be a reasonable approximation for the boundary between classes even in the presence of homoploutia. We thank Malcolm Sawyer for pointing us to this argument.

- Alpizar, F., Carlsson, F., and Johansson-Stenman, O. (2005). How much do we care about absolute versus relative income and consumption? *Journal of Economic Behavior & Organization*, 56(3):405–421.
- Ascione, F. and Schnetzer, M. (2022). Out of balance? Revisiting the nexus of income inequality, household debt and current account imbalances after the Great Recession. *Cambridge Journal of Economics*, 46(1):133–165.
- Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2(3):244–263.
- Atkinson, A. B. (2009). Factor shares: the principal problem of political economy? *Oxford Review of Economic Policy*, 25(1):3–16.
- Barbosa-Filho, N. H. and Taylor, L. (2006). Distributive and demand cycles in the US economy – a structuralist Goodwin model. *Metroeconomica*, 57(3):389–411.
- Battistin, E., Blundell, R., and Lewbel, A. (2009). Why is consumption more log normal than income? Gibrat’s law revisited. *Journal of Political Economy*, 117(6):1140–1154.
- Behringer, J. and van Treeck, T. (2018). Varieties of capitalism and growth regimes: The role of income distribution. *IMK Working Paper*, 193.
- Berg, A., Ostry, J. D., Tsangarides, C. G., and Yakhshilikov, Y. (2018). Redistribution, inequality, and growth: new evidence. *Journal of Economic Growth*, 23(3):259–305.
- Berman, Y. and Milanović, B. (2020). Homoploutia: Top labor and capital incomes in the United States, 1950 – 2020. *Stone Center on Economic Inequality Working Paper Series*, 28.
- Bertrand, M. and Morse, A. (2016). Trickle-down consumption. *Review of Economics and Statistics*, 98(5):863–879.
- Bhaduri, A. and Marglin, S. (1990). Unemployment and the real wage: the economic basis for contesting political ideologies. *Cambridge Journal of Economics*, 14(4):375–393.
- Bricker, J., Krimmel, J., and Ramcharan, R. (2020). Signaling status: The impact of relative income on household consumption and financial decisions. *Management Science*, 67(4):1–17.
- Brooks, D. (2010). *Bobos in paradise: The new upper class and how they got there*. Simon and Schuster.
- Brzozowski, M., Gervais, M., Klein, P., and Suzuki, M. (2010). Consumption, income, and wealth inequality in Canada. *Review of Economic Dynamics*, 13(1):52–75.
- Chakrabarti, A. S., Chatterjee, A., Nandi, T., Ghosh, A., and Chakraborti, A. (2018). Quantifying invariant features of within-group inequality in consumption across groups. *Journal of Economic Interaction and Coordination*, 13(3):469–490.
- Charles, K. K., Hurst, E., and Roussanov, N. (2009). Conspicuous consumption and race. *Quarterly Journal of Economics*, 124(2):425–467.

- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022a). Social capital II: determinants of economic connectedness. *Nature*, 608:122–134.
- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhole, M., and Wernerfelt, N. (2022b). Social capital I: Measurement and Associations with Economic Mobility. *Nature*, 608:108 – 121.
- Clementi, F. and Gianmoena, L. (2017). Modeling the joint distribution of income and consumption in Italy: A copula-based approach with  $\kappa$ -generalized margins. In Galletti, M., Palestini, A., and Russo, A., editors, *Introduction to Agent-Based Economics*, pages 191–228. Elsevier.
- Colander, D. C. (1993). The macrofoundations of micro. *Eastern Economic Journal*, 19(4):447–457.
- Davis, F. (1992). *Fashion, culture, and identity*. The University of Chicago Press.
- Dence, T. P. and Dence, J. B. (2009). A survey of Euler’s constant. *Mathematics Magazine*, 82(4):255–265.
- Detzer, D. (2018). Inequality, emulation and debt: The occurrence of different growth regimes in the age of financialization in a stock-flow consistent model. *Journal of Post Keynesian Economics*, 41(2):284–315.
- Drechsel-Grau, M. and Schmid, K. D. (2014). Consumption–savings decisions under upward-looking comparisons. *Journal of Economic Behavior & Organization*, 106:254–268.
- Dynan, K. E., Skinner, J., and Zeldes, S. P. (2004). Do the rich save more? *Journal of Political Economy*, 112(2):397–444.
- Eichner, A. S. (1986). *Toward a new economics: Essays in post-Keynesian and institutionalist theory*. MacMillan.
- Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5(1):17–60.
- Fagiolo, G., Alessi, L., Barigozzi, M., and Capasso, M. (2010). On the distributional properties of household consumption expenditures: The case of Italy. *Empirical Economics*, 38(3):717–741.
- Fagiolo, G., Guerini, M., Lamperti, F., Moneta, A., and Roventini, A. (2019). Validation of agent-based models in economics and finance. In Beisbart, C. and Saam, N. J., editors, *Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives*, pages 763–787. Springer International Publishing, Cham.
- Flechtner, S. and Gräbner, C. (2019). The heterogeneous relationship between income and inequality: a panel co-integration approach. *Economics Bulletin*, 39(4):2540–2549.

- Frank, R., Levine, A., and Dijk, O. (2014). Expenditure Cascades. *Review of Behavioral Economics*, 1:55–73.
- Friedman, S. and Reeves, A. (2020). From aristocratic to ordinary: Shifting modes of elite distinction. *American Sociological Review*, 85(2):323–350.
- Gabaix, X. and Landier, A. (2008). Why has CEO pay increased so much? *The Quarterly Journal of Economics*, 123(1):49–100.
- Ger, G. and Belk, R. W. (1999). Accounting for materialism in four cultures. *Journal of Material Culture*, 4(2):183–204.
- GESIS – Leibniz-Institut für Sozialwissenschaften (2011). ALLBUS/GGSS 2006 (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften/German General Social Survey 2006). GESIS Data Archive, Cologne. ZA4500 Data file Version 2.0.0, <https://doi.org/10.4232/1.10832>.
- Ghosh, A., Gangopadhyay, K., and Basu, B. (2011). Consumer expenditure distribution in India, 1983–2007: Evidence of a long Pareto tail. *Physica A: Statistical Mechanics and its Applications*, 390(1):83–97.
- Gladys Foster, P. (1990). Keynes and Kalecki on saving and profit: Some implications. *Journal of Economic Issues*, 24(2):415–422.
- Gräbner, C. (2018). How to relate models to reality? An epistemological framework for the validation and verification of computational models. *Journal of Artificial Societies and Social Simulation*, 21(3).
- Harris, D. J. (1974). The price policy of firms, the level of employment and distribution of income in the short run. *Australian Economic Papers*, 13(22):144–151.
- Heffetz, O. (2011). A test of conspicuous consumption: Visibility and income elasticities. *Review of Economics and Statistics*, 93(4):1101–1117.
- Hein, E. (2014). *Distribution and growth after Keynes: A Post-Keynesian guide*. Edward Elgar Publishing.
- Hohnisch, M., Pittnauer, S., and Chakrabarty, M. (2002). Empirical regularities in distributions of individual consumption expenditure. *International Journal of Modern Physics C*, 13(04):541–549.
- Iacono, R. and Ranaldi, M. (2021). The evolution of income composition inequality in Italy, 1989 – 2016. *Review of Income and Wealth*, forthcoming.
- Izyumov, A. and Vahaly, J. (2015). Income shares revisited. *Review of Income and Wealth*, 61(1):179–188.
- Jappelli, T. and Pistaferri, L. (2010). Does consumption inequality track income inequality in Italy? *Review of Economic Dynamics*, 13(1):133–153.
- Jenkins, D. (2016). Conspicuous consumption in the United States and China. *Journal of Economic Behavior & Organization*, 127:115–132.



- Johnson, D. G. (1954). The functional distribution of income in the United States, 1850-1952. *Review of Economics and Statistics*, 35(2):175–182.
- Kalecki, M. (1971). *Selected essays on the dynamics of the capitalist economy 1933-1970*. Cambridge University Press.
- Kapeller, J. and Schütz, B. (2015). Conspicuous consumption, inequality and debt: the nature of consumption-driven profit-led regimes. *Metroeconomica*, 66(1):51–70.
- Kharroubi, E. and Kohlscheen, E. (2017). Consumption-led expansions. *BIS Quarterly Review*, March.
- Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives*, 6(2):117–136.
- Kossinets, G. and Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, 115(2):405–450.
- Krueger, D. and Perri, F. (2006). Does income inequality lead to consumption inequality? Evidence and theory. *Review of Economic Studies*, 73(1):163–193.
- Kumar, R. (2017). 'Two Class' Distribution of Income in India. *Economic & Political Weekly*, 52(6):59.
- Kumar, R. (2021). Personal income inequality in USA from a two-class perspective: 2004-2018. *SocArXiv*, fmkj3.
- Lavoie, M. (1992). *Foundations of Post-Keynesian Economic Analysis*. Edward Elgar Publishing.
- Lavoie, M. (1994). A post Keynesian approach to consumer choice. *Journal of Post Keynesian Economics*, 16(4):539–562.
- Lavoie, M. (2022). *Post-Keynesian economics: new foundations*. Edward Elgar Publishing, 2 edition.
- Mac Carron, P., Kaski, K., and Dunbar, R. (2016). Calling Dunbar's numbers. *Social Networks*, 47:151 – 155.
- Mayerhoffer, D. M. and Schulz, J. (2021). A Network Approach to Consumption. *BERG Working Paper Series*, 163.
- Mayerhoffer, D. M. and Schulz, J. (2022). Perception and privilege. *Applied Network Science*, 7(1):1–25.
- McPherson, J. M. and Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review*, 52(3):370–379.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.
- Milanovic, B. (2005). *Worlds Apart: Measuring International and Global Inequality*. Princeton University Press.

- Mott, T. and Slatiery, E. (1994). Tax incidence and macroeconomic effects in a Kaleckian model when profits finance investment and prices may respond to taxes. *Journal of Post Keynesian Economics*, 16(3):391–410.
- Onaran, O. and Obst, T. (2016). Wage-led growth in the EU15 member-states: the effects of income distribution on growth, investment, trade balance and inflation. *Cambridge Journal of Economics*, 40(6):1517–1551.
- Peters, O. (2019). The ergodicity problem in economics. *Nature Physics*, 15(12):1216–1221.
- Proaño, C. R., Flaschel, P., Krolzig, H.-M., and Diallo, M. B. (2011). Monetary policy and macroeconomic stability under alternative demand regimes. *Cambridge Journal of Economics*, 35(3):569–585.
- Quintana-Domeque, C. and Wohlfart, J. (2016). “Relative concerns for consumption at the top”: An intertemporal analysis for the UK. *Journal of Economic Behavior & Organization*, 129:172–194.
- Reich, M., Gordon, D. M., and Edwards, R. C. (1973). A theory of labor market segmentation. *American Economic Review*, 63(2):359–365.
- Rényi, A. (1953). On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Hungarica*, 4(2).
- Robinson, J. (1956). *The accumulation of capital*. MacMillan.
- Saez, E. and Zucman, G. (2019). *The triumph of injustice: How the rich dodge taxes and how to make them pay*. WW Norton & Company.
- Schneider, M. (2015). Revisiting the thermal and superthermal two-class distribution of incomes. *European Physical Journal B*, 88(1):1–10.
- Schneider, M. and Scharfenaker, E. (2020). Mixing it up. *European Physical Journal Special Topics*, 229(9):1685–1704.
- Schulz, J. and Mayerhoffer, D. M. (2022). A network approach to consumption. *arXiv preprint*, 2203.14259.
- Schulz, J., Mayerhoffer, D. M., and Gebhard, A. (2022). A Network-Based Theory of Inequality Perceptions. *Social Networks*, 70.
- Setterfield, M. and Kim, Y. K. (2017). Household borrowing and the possibility of ‘consumption-driven, profit-led growth’. *Review of Keynesian Economics*, 5(1):43–60.
- Shaikh, A. (2017). Income distribution, econophysics and Piketty. *Review of Political Economy*, 29(1):18–29.
- Shaikh, A. and Jacobo, J. E. (2020). Economic arbitrage and the econophysics of income inequality. *Review of Behavioral Economics*, 7:299–315.
- Sherman, R. (2018). ‘A very expensive ordinary life’: consumption, symbolic boundaries and moral legitimacy among New York elites. *Socio-Economic Review*, 16(2):411–433.

- Shiller, R. J. (2019). *Narrative economics*. Princeton University Press Princeton.
- Silva, A. C. and Yakovenko, V. M. (2004). Temporal evolution of the “thermal” and “superthermal” income classes in the USA during 1983–2001. *Europhysics Letters*, 69(2):304.
- Simon, H. A. (1976). From substantive to procedural rationality. In Kastelein, T. J., Kuipers, S. K., Nijenhuis, W. A., and Wagenaar, G. R., editors, *25 years of economic theory*, pages 65–86. Springer.
- Stanley, T. J. and Danko, W. D. (1996). *The millionaire next door: The surprising secrets of America’s wealthy*. Taylor Trade Publishing.
- Stockhammer, E. (2017). Wage-led versus profit-led demand: what have we learned? A Kaleckian–Minskyan view. *Review of Keynesian Economics*, 5(1):25–42.
- Stockhammer, E. and Stehrer, R. (2011). Goodwin or Kalecki in demand? Functional income distribution and aggregate demand in the short run. *Review of Radical Political Economics*, 43(4):506–522.
- Talaga, S. and Nowak, A. (2020). Homophily as a Process Generating Social Networks: Insights from Social Distance Attachment Model. *Journal of Artificial Societies and Social Simulation*, 23(2):6.
- Toda, A. A. (2017). A note on the size distribution of consumption: More double Pareto than lognormal. *Macroeconomic Dynamics*, 21(6):1508–1518.
- Wang, J.-L., Wang, H.-Z., Gaskin, J., and Hawk, S. (2017). The mediating roles of upward social comparison and self-esteem and the moderating role of social comparison orientation in the association between social networking site usage and subjective well-being. *Frontiers in Psychology*, 8(771):1–9.
- World Bank (2022). World development indicators. <https://data.worldbank.org/indicator/NE.CON.PRVT.ZS>. [Online; accessed 2022-07-17].
- Yanev, G. P. (2020). Exponential and hypoexponential distributions: Some characterizations. *Mathematics*, 8(12):2207.