Essays on the Economics of Heterogeneity

A thesis submitted in partial fulfillment of the requirements

of the degree of Doctor of Philosophy (Ph.D.) in Economics

Nils-Holger Gudat

School of Economics and Finance

Queen Mary, University of London

October 2015

To Nadien, who lifted my spirits more than once over the last four years
and was more understanding and patient than anyone could hope for;
my mother, as I couldn't have made it here without her love and support;
and Emil, whose laugh carried me through the final stages of this project.

OMAHA!

Declaration

I wish to declare

I, Nils-Holger Gudat, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Details of collaboration and publications:
Parts of Chapter 5 were undertaken as joint work with Ryan Weldzius (University of California, Los Angeles).

Signature: Nils-Holger Gudat
Date: October 2, 2015

Acknowledgements

I would like to thank my supervisors, Xavier Mateos-Planas and Winfried Koenieger, for their guidance, patience and support throughout my Ph.D. studies.

I further acknowledge the School of Economics and Finance at Queen Mary, University of London, for generous funding.

Abstract

The analysis of the effects of heterogeneity on aggregate economic outcomes has seen a resurgence in the recent macroeconomic literature. The exponential increase in computer power over the last decades has allowed researchers to solve ever more complex theoretical models with meaningful heterogeneity along various dimensions, while at the same time bringing ever more granular micro-level data to the table when testing the model predictions.

This thesis explores two varieties of this recent vintage of models of heterogeneity. The first part of the thesis investigates the implications for wealth distributions of combining the standard life-cycle incomplete markets model of household consumption with income processes featuring heterogeneity in individual-specific growth rates, which households can learn about over the course of their working life. To this extent, first the recent literature on partial insurance and models of wealth inequality is reviewed. Then, income processes with profile heterogeneity are estimated from PSID and BHPS data. The results confirm the findings of previous studies that allowing for profile heterogeneity significantly lowers the estimated persistence and innovation variance of persistent shocks to household income, and documents substantial variation in the

estimated parameters of the income process across time periods and measures of household income. The estimated income processes obtained are then used in a quantitative model of household consumption and saving in order to investigate the implications for the model predictions on the wealth distribution. The model is calibrated to empirical wealth distributions obtained from the SCF and the BHPS, and it is shown that the inclusion of individual-specific growth rate heterogeneity in income severely deteriorates the models ability to fit the shape of the data. Comparative statics exercises are performed to identify the drivers in the models failure to match the empirical profile of wealth holdings, which show that it is precisely the two key parameters which differ between the standard AR(1) model and the heterogeneous profile model, the persistence and variance of the permanent shock to household income, which drive model fit. The second part of the thesis looks at heterogeneity on the production side of the economy and its implications for international trade. Following an existing approach in the literature, we develop testable implications of the Melitz and Ottaviano (2008) model of trade, in which firms differ in their productivity and have to make production and exporting decisions in the face of costs to trade. Applying an estimation strategy previously used in the literature, we find weak support of the models predictions in data for 64 manufacturing industries in the NAFTA member countries Canada, Mexico and USA. We then test additional model predictions by constructing a measure of entry conditions by industry based on firm turnover, which allows us to divide our sample into fixed and free entry industries. Furthermore, we include the effects of third country tariff barriers on the relative performance of two trading partners' industries. While the results are broadly in line with model predictions, we find some evidence of violations of the predictions in the data.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Distributional questions are increasingly making a comeback in economics. In spite of the famous – or infamous – warning of Lucas (2004) that the focus on questions of distribution is one of the most ”seductive (...) and poisoning” tendencies in economics[1], many fields of economics that have long relied on simplistic models of representative households and firms have increasingly taken the issue of modelling heterogeneity across economic agents seriously. At the very least since the Great Recession triggered by the financial crisis of 2007–2008, issues of distribution have also taken centre stage in the public economic discourse. Work on the increase in income inequality, especially at the top end of the income distribution, and the rising inequality in wealth holdings in advanced

[1] In fairness it has to be said that Lucas' quote is often taken out of context, as he was not actually advising against studying distributional issues entirely, but merely pointing out that economic growth has played a much more important role in raising people out of poverty than re- distribution of current resources at any point in time could have achieved.

economies has played a prominent role in the public debate in recent years, the most prominent recent example being Piketty (2014), a rare instance of a book largely based on economic scholarship being widely discussed and sold (if maybe not read) by a mainstream audience. However, while the public has only recently started to take an interest in issues of inequality and distribution, the economic literature has been developing quantitative models of heterogeneity for almost three decades. Seminal papers such as Imrohoruglu (1989), Huggett (1993), and Aiyagari (1994) have laid the groundwork for a vast literature explicitly modelling the choices of heterogeneous agents based on microeconomic evidence. A major factor in the move towards models of explicit heterogeneity have been the huge advancements in computer power in recent decades. With Moore's law still holding to this day, the transistor count of the fastest microprocessor today is about two-thousand times as high as that of the fastest microprocessor twenty-five years ago, when Zeldes (1989) published one of the first works that numerically solved a household savings problem under uncertainty numerically. Besides enabling researchers to numerically solve ever more complex optimisation problems with state spaces of ever more dimensions, the advancements in computer power have also vastly improved data processing capabilities, a development that in turn has led to a surge in empirical work exploiting large microeconomic data sets, the results of which can then be used to validate models of household and firm behaviour.

The present work explores heterogeneity in two different classes of economic models, using both approaches outlined above. While the first part will build up towards a quantitative theoretical model of the wealth distribution, the merits of which will be evaluated against micro survey data, the second part will take the predictions of a micro-founded model of international trade and test them on data of firm behaviour disaggregated at the industry-level.

1.2 Outline

This thesis is structured as follows. Chapter 2 gives an overview of the research on theoretical models of the household wealth distribution in the last two decades. It will highlight the key empirical challenges by presenting stylised facts on the cross sectional wealth distribution using the most recent wave of the UK Wealth and Asset Survey as an example. Then, the workhorse model of the literature on household savings decision, the incomplete markets life-cycle model of consumption, is briefly reviewed along with recent work partial insurance in this model. Then, extensions of the model which help it match the stylised facts of empirical wealth distributions are reviewed.

Chapter 3 builds on the discussion in Chapter 2 by estimating income processes with profile heterogeneity for different sub-periods of PSID data from 1968 to 2013 and from BHPS data in order to assess the stability of the cross-sectional variance of income growth rates across time and income measures. The estimates are then used in chapter 4 as inputs in a structural model of household saving first employed by Guvenen (2007), which is calibrated to the empirical wealth distribution using a minimum distance estimator. After discussing the model fit, comparative statics exercises are performed on all parameters of the income process, the upper and lower bounds of which are taken from the universe of estimation results from chapter 3, to understand which parameters are most important for the model fit.

Chapter 5 gives a brief introduction into trade models based on firm-level heterogeneity, before developing an estimable model in the spirit of Chen et al. (2009). The model is then tested on a data set of prices, productivity, and markups for 64 manufacturing industries in the Canada, Mexico and the United States over the period of 1988 to 2010. In an extension of the approach of Chen et al. (2009), a sub-sample analysis is conducted in which the observations are split into fixed and free entry industries, based on a measure of firm turnover developed on the basis of prior research.

Chapter 5 concludes the thesis and discusses potential avenues for future research.

Chapter 2

Quantitative Models of the Wealth Distribution

2.1 Introduction

The distribution of wealth and income has recently made a comeback to the centre of economic discourse in advanced economies. The ongoing rise of income inequality, observed since the early 1980s especially in Anglo-Saxon countries, has received renewed attention in the public sphere since the financial crisis started taking its toll on living standards across the world. At the same time, the best-selling book by Piketty (2014) led to a surge in interest in the role of capital in the economy, and, by extension, the distribution of wealth, both in the academic literature and the popular press.

While the broader public has only recently picked up on the issues arising around income and wealth distributions, they have sat squarely in the centre of many sub-fields of economics for a long time. The income distribution has long been of interest to labour economists trying to understand the forces shaping the evolution of earnings in the labour market, while at least the accumulation of aggregate wealth plays a central role in macroeconomic models of economic growth. This chapter, as well as chapter 4, focuses on the intermediate step that takes us from an income to a wealth distribution - economic models of household saving. When attempting to build a model of the wealth distribution, the first step of course is to get an understanding of the object we want to model. To this end, this chapter starts by presenting stylised facts of the wealth distributions in advanced countries and discusses some of the limitations of the data available. It then builds a simple life-cycle model of consumption and savings to guide the following discussion and fix notation. Using this basic model, different savings motives and their importance in the context of aggregate wealth accumulation are discussed. Following this, the role of income uncertainty and market structure is examined in more detail.

2.2 Stylised facts

The most notable and consistent fact that emerges when looking at wealth distributions across all countries and different time periods is that wealth is highly unevenly distributed, much more so than income. Vermeulen (2014) argues that wealth holdings are so concentrated, that even surveys employing designs that feature oversampling of richer households (such as the US Survey of Consumer Finances (SCF), or the British Wealth and Asset Survey (WAS)) underestimate the percentage share of wealth held by the highest percentile of the wealth distribution by anywhere from one to five percentage points, while surveys that don't oversample underestimate the share by up to ten percentage points. Keeping this in mind, figure 2.1 presents evidence on household wealth from the most recent wave of the WAS, which surveyed over 20,000 British households in the years 2010 to 2012. The first thing that becomes obvious from the figure is the very unequal distribution of wealth. While households in the lowest two percentiles hold virtually no wealth apart from physical wealth, which includes possessions such as vehicles and furniture, wealth holdings grow exponentially when we move up the distribution, with the top percentile holding twice as much wealth as the ninth percentile. It is important to note that in producing these aggregate numbers and figures, one necessarily has to make decisions on how exactly to construct and aggregate measures of wealth, which will have to be kept in mind when comparing model predictions with empirical numbers. If the goal is to account for all productive capital in the economy that can be used in production, a measure of total net wealth aggregating all forms of asset and debt classes, and including some durable consumption goods such as cars. When thinking about the role of wealth in helping the household to smooth out income fluctuations, it might be more appropriate to exclude very illiquid assets such as housing, and look more closely at the role of debt for households which might be at their borrowing constraint and are thus vulnerable to reductions in their borrowing limit, even though their net wealth (including illiquid assets) is positive. Finally, important questions are raised by the existence of various government and private pension schemes, which have to be factored in when constructing measures of a household's lifetime resources, but whose exact value might be uncertain (for the case of defined contribution plans) and not well understood by households themselves. Figure 2.1 makes clear that this is a non-trivial issue, as implied

Figure 2.1: Histograms and cdfs of household net wealth in different data sets.

pension wealth is a significant portion of household portfolios.

2.3 A workhorse model

The basic model underlying the discussion of savings behaviour and wealth accumulation in this chapter is the life-cycle model of household behaviour dating back to Modigliano and Brumberg (1954) [2]. The model can be written as a single household solving the problem

[2]A more detailed treatment of the general class of models can be found both in Browning and Crossley (2001) and in Attanasio and Weber (2010), although both papers have their focus on household consumption behaviour rather than wealth accumulation.

$$\max_{\{c_{t+j}\}_{j=0}^{T-t}} \sum_{j=0}^{\infty} \delta^{t+j} \mathbb{E}_t \big[u(c_{t+j}, z_{t+j}) \big] \tag{2.1}$$

subject to

$$a_{t+1} = (1+r)(a_t + y_t - c_t) \tag{2.2}$$

where T is the last period of the planning horizon, δ is the subjective discount factor, $u()$ is the instantaneous felicity function – usually assumed to be of the CRRA form, $\frac{c^{1-\sigma}}{1-\sigma}$, where σ is the coefficient of relative risk aversion –, c is consumption, a are financial assets, which allow the household to transfer resources across time, r is a one-period interest rate, and y is income. z is used as a stand-in variable denoting the fact that households might, in general, care about other things that are not captured by the concept of current consumption; examples would be habits or durable consumption (which break the time-separability of the utility function), leisure time, particular classes of assets (such as housing) or bequests left to future generations. While in general, $T \to \infty$ is a possibility, and infinite horizon versions – first advocated by Friedman (1957) – of the life-cycle model are widely used in macroeconomic applications, the finite horizon model will be more useful for the following discussion and forms the centrepiece of this thesis for a number of reasons which will become clear as we progress.

2.4 Saving motives

When trying to understand wealth distributions through the lens of the model outlined above, the key question is: why do households save? The basic model, in which consumers only care about the time path of instantaneous utility derived from consumption, suggests that households will save if and only if it leads to preferential allocation of consumption over time. If the utility function exhibits curvature, as the standard CRRA utility function does, households will prefer a smooth consumption path over time. As Browning and Crossley (2001) point out, consumption smoothing can happen at different frequencies, depending on the exact set-up of the model. In Modigliano and Brumberg (1954) the main reason saving was the existence of a retirement period, which necessitates consumption smoothing over the life cycle – wealth accumulation during working life to pay for consumption in retirement. The implication of this simple model is that wealth accumulation on the household level solely depends on the length of the retirement period, while in aggregate wealth accumulation crucially hinges on the growth rate of the economy. The crucial assumption that allows Modigliani and Brumberg to focus on consumption smoothing over the life cycle was that of constant income, an assumption that is obviously incorrect and easily rejected by the data. We will therefore next discuss the implications of introducing income fluctuations, which will make consumption smoothing at shorter frequencies necessary, into the basic model.

2.5 Income uncertainty and market structure

With the assumption of a non-constant income stream, it becomes important to think about the opportunities households have to insure themselves against these fluctuations, or, in other words, which market structure they are facing. To make income fluctuations relevant for the economic agent, the world of complete markets, in which a full set of Arrow securities covering each possible state of the world can be bought and sold, has to be abandoned in favour of market *incompleteness*. The most convenient, and at the same time most extreme, departure from the complete markets assumption is to assume away any sort of insurance markets except for very simple self-insurance through risk- free one-period bonds. This market structure is implicit in the formulation of the consumer problem in equation (2.2) – there is just one asset for the household to sell or buy, and this asset has a certain payoff in the following period, which is not contingent on the state of the world. The big advantage of this setup is tractability: simple models of this kind can often be solved analytically, and in recursive formulations of more complex problems, the simple market structure only adds one state variable to the problem. The drawback, obviously, is that this market structure is at odds with the economic reality, where households are able to buy a hoist of different assets that vary widely in liquidity as well as in the degree to which payoff are state-contingent. We defer the consideration of the role of liquidity to section 2.6, which deals with the largest asset in most households' portfolio, housing, and examine the role of insurance first. The basic idea when investigating the extent to which households have access to insurance mechanisms is to analyse the joint dynamics of income and consumption data, and compare them with the implications derived from models with different insurance mechanisms. In a complete market setup, where households can fully insure income risk, idiosyncratic changes in income should not translate into changes in consumption, implying a flat profile of cross-sectional consumption inequality over the life- cycle, irrespective of the underlying stochastic process governing income. This is not the case in the absence of insurance opportunities, with the opposite end of the model spectrum being inhabited by the Aiyagari-Bewley-Hugget-Imrohoroglu class of models (Aiyagari (1994), Bewley (1977), Huggett (1993), Imrohoruglu (1989)). These model don't feature any insurance possibilities apart from non-contingent one-period bonds and also deliver specific predictions on the relationship between income, consumption and savings[3]. The precise predictions of the model for how households will consume and save depend crucially on the specification of the stochastic process governing income uncertainty – essentially the object over which $\mathbb{E}$ in equation (4.1) is defined. When applying an income process consisting of permanent and transitory shocks to this model, the well-known[4] result of the model is that consumption should react to permanent changes in income, while transitory changes in income should be buffered by saving and dissaving in the noncontingent bond. This prediction of the model is exploited by some authors to elicit information on the decomposition of income changes into permanent and transitory shocks using consumption data: assuming that the model is correct, increases in income inequality in the data that are accompanied by contemporary increases in consumption inequality must be

[3]More precisely, the opposite end of the insurance spectrum would be a world that does not even offer noncontingent bonds, although this market structure is obviously not suited to examine any interesting economic question.

[4]For a rigorous derivation refer to your favourite Macroeconomics textbook, e.g. Ljungqvist and Sargent (2012), chapter 17.

induced by permanent shocks to income, while changes in the income distribution that do not lead to changes in the consumption distribution can be seen to be the consequence of transitory shocks. Blundell and Preston (1998) is an example of a paper employing exactly this strategy to examine data on consumption and income from the British Family Expenditure Survey to examine the properties of changes in the income distribution in Britain between 1968 and 1992. One problem of these studies however is a large literature documenting ”excess smoothness” of consumption in the data, that is showing that consumption does *not* change one-for-one even with changes in income that are known to be permanent (are very detailed account on the early research on this can be found in Deaton 1992), implying that there are at least some insurance opportunities available to households in the real world. Based on the rejection of both full and no insurance in the data, an active literature has developed trying to quantify the amount of insurance households have access to. Blundell et al. (2008) develop a novel imputation procedure designed to alleviate measurement problems in PSID consumption data to test the degree of partial insurance. Their estimates imply that households are almost perfectly able to insulate consumption from transitory shocks, as a standard Bewley model would predict, but, in contrast to the model predictions, also that around 40% of permanent shocks to income can be insured against. Kaplan and Violante (2010) examine to what extent the empirical estimates of consumption insurance that Blundell et al. (2008) obtain can be replicated in a standard incomplete market model with capital as the only savings vehicle. They find that the model replicates the high degree of insurance against transitory shocks, but fails to generate enough insurance against permanent shock;

a version of the model with a natural borrowing constraint[5] generates an insurance coefficient of 22%, while one without borrowing only delivers 7% insurance. They also show that the model generates too much insurance for older workers and too little for younger workers, which, just as the importance of the role of borrowing constraints, is consistent with the empirical finding in Blundell et al. (2008) that insurance coefficients are much higher in a sample of high wealth households compared to a sample of low wealth households. All of this suggests that in order to accurately assess self insurance in a quantitative model, it is important for the model to capture the wealth accumulation adequately, as wealth holdings might significantly change the insurance options of households, even if that wealth was accumulated for reasons entirely different from the self-insurance motives present in the simple incomplete markets model.

2.6 The role of housing

One important component of household saving for which the motives to accumulate it are more complex than a simple self-insurance model would imply is housing. As figure 2.1 demonstrates using data from the most recent British Wealth and Asset Survey, by far the largest share of household portfolios is invested in housing wealth, with the notable exception of the very richest households. Similar portfolio allocations can be observed for all advanced economies, Fernandez-Villaverde and Krueger (2011) for example show that households between the 30th and 80th percentile of the wealth distribution

[5]This is one of the loosest possible borrowing constraints commonly used in the literature, it implies that households can borrow up to the capitalised value of the income stream until the end of life that would obtain if they were hit by the worst sequence of income shocks possible.

hold in excess of 60% of their total assets in housing wealth. The illiquidity of housing, the high transaction costs and the consumption element of housing purchases make this asset fundamentally different from the one-period riskless bond considered in our workhorse model, and imply that it offers only very limited insurance opportunities against short-term income fluctuations. Furthermore, housing differs from other assets in that it also offers the household housing services, that is an investment in real estate is at the same time the purchase of a durable good. Hence, the introduction of housing into our simple model above does not only mean a departure from having a simple asset a_t in equation (2.2), it also means having an additional consumption good which does not perish, possibly altering the households utility function in equation (2.1), if the utility derived from the housing good is of a different form that that derived from instantaneous consumption. A number of authors have considered the effects of allowing households to save in housing assets in addition to financial assets. Yang (2009) constructs synthetic cohorts to examine consumption data from the CEX in tandem with asset allocation data from the SCF and documents diverging consumption patterns over the life-cycle for housing compared to non-housing goods. She also shows that wealth accumulation is markedly different for home owners compared to renters, with home owners holding much more wealth in retirement than non-owners, in contrast to the predictions of a simple life-cycle model. She also shows that a model which includes housing services in the utility function and features borrowing constraints coupled with downpayments for housing purchases (which prevent many agents from accumulating housing wealth early in life) and transaction costs in the adjustment of the housing stock (which slows the decumulation of housing assets at the end of life) can match the key empirical facts. Fernandez-Villaverde and Krueger (2011) present an argument along similar lines and build a model with endogenous borrowing constraints in which durable consumption goods (which can be seen as housing) act as collateral. Here, the model predicts accumulation of durable goods early in life, with high consumption of nondurables and accumulation of financial assets later in life, in line with the data. Iacoviello (2008) considers a similar model and shows that the presence of housing assets in the model can lead to a decoupling of the joint evolution of cross-sectional consumption and income inequality that was at the heart of the literature identifying permanent and transitory income shocks from precisely this relationship. This again highlights the crucial importance of getting the mechanisms driving households' wealth accumulation right if one wants to draw conclusions about the risk households are facing from quantitative models of household consumption and savings.

2.7 Closed and open economies

One key decision when building a model of wealth accumulation is the question of how the interest rate on savings is determined. Traditionally, the macroeconomic literature has viewed the interest rate as an endogenous parameter, pinned down by the marginal product of capital from the economy's production function and the quantity of capital available, which in turn is governed by household's savings decisions. In fact, the main contribution of the seminal work by Aiyagari (1994) was to highlight the effect of idiosyncratic income risk and borrowing constraints, two key features of the heterogeneous agent models most frequently used to examine questions related to the wealth distribution, on the steady state interest rate. Aiyagari shows that compared to a standard growth model, the steady state stock of capital in a closed incomplete markets economy is higher, and, correspondingly, the steady state interest rate is lower because of precautionary savings induced by income uncertainty. However, allowing for an endogenous interest rates introduces additional complexity into the model as the determination of the interest rate requires asset market clearing, which implies that household savings choices have to be consistent with each other in each period. For many modern life-cycle models which feature large, high-dimensional state spaces, this additional computational burden might make the model solution infeasible. For these reasons, many researchers have opted for treating the interest rate as an exogenous parameter, set anywhere between three (Cagetti, 2003) and 5.2 (Gourinchas and Parker, 2002) percent. Whether this is justified will depend on two things: theoretically, one has to ask whether general equilibrium effects are likely to alter the answer to the question at hand, while empirically the question of how high the elasticity of the interest rate to changes in aggregate wealth is will determine how problematic omitting this feedback mechanism from the model is. In general equilibrium models in the tradition of Aiyagari (1994), this elasticity is given by the sensitivity of the marginal product of capital to the quantity of aggregate capital, which can be determined from the production function. While the Cobb-Douglas function is the function of choice in the literature (see e.g. Castaneda et al. 2003), recent work by Piketty (2014) casts doubt on the appropriateness of this assumption and argues for a functional form implying a lower elasticity of the interest rate to increases in the capital stock. Irrespective of the choice of the production function though, one has to ask how valid the assumption of a closed economy, in which household savings have direct impacts on the quantity of productive capital in the economy, is. Given the deep international integration of modern financial markets, it appears that the open economy assumption often used in international economics to describe economies that can not set interest rates might be useful when thinking about the dynamics of interest rates in response to changes in saving behaviour in the local economy. Davies et al. (2007) show that even the US as the economy with the – by far – largest net wealth holdings only accounts for around 30% of global net wealth. This suggests that any excess negative or positive asset holding in an economy for which the asset market clearing condition does not hold locally can be absorbed by global financial markets. Indeed, Bernanke (2005) famously argued that a ”global savings glut” was sustaining large US current account deficits, suggesting that global financial markets allow even the largest asset market in the world to not clear for extended periods of time. These considerations motivate us to opt for an exogenous interest rate in chapter 4.

2.8 Wealth Distribution Papers

In the last years, many authors have used the possibilities offered by the increase in computing power to derive an additional implication from the broad class of incomplete market models outlined above: a simulated wealth distribution. While there are many practical difficulties in creating model outputs that can reasonably be compared to the data collected in surveys (some of which have been alluded to in the above discussion on the definition of wealth), in principle the simulated wealth distribution derived from life-cycle models can be used to calibrate deep parameters of the model, provide an additional test for how well the model is able to capture household savings behaviour, and shed light on which mechanisms are crucial in driving the evolution of aggregate savings at different parts of the distribution.

An early attempt to use the wealth distribution to estimate the parameters of a life-cycle model of household savings can be found in Cagetti (2003), who uses a simple model similar to the one outlined in equations (2.1) and (2.2). Important additions in his version of the model are a bequest motive – which, ceteris paribus, will increase the wealth holdings of elderly households – and a simplified pension system, which guarantees each household a pension depending on their education level, and thereby lowers wealth accumulation during working life. The idea behind the estimation strategy is simple: given a stochastic process for household income, the discount rate δ and the risk aversion parameter σ pin down a solution to the household's savings problem which allows one to simulate a theoretical wealth distribution from optimal household behaviour. Therefore, it is possible to use the simulated method of moments to construct an estimator that chooses the vector (δ, σ) which minimises the distance between empirical moments of the wealth distribution and its simulated counterparts. Given the high skewness of wealth data, Cagetti opts for median wealth by 5-year age group as the moment to match. As has become clear in the previous discussion, a crucial element driving household choices in the model is the income risk they face, making the choice for the stochastic process representing this risk and its calibration a crucial step in modelling wealth distributions. Cagetti opts for a process consisting of a trend growth component common to all households, an age-education component estimated for CEX data, and an MA(1) process representing the stochastic nature of income. With his calibration, Cagetti finds low degrees of persistence, with pronounced heterogeneity across education groups, and high degrees of risk aversion, implying a significant contribution of precautionary savings to aggregate wealth.

A very similar exercise is performed by Hintermaier and Koeniger (2011), who construct a minimum distance estimator based on the shape of the cross-sectional distribution of wealth at different stages of the life cycle. That is, rather than simply targeting the 50th percentile of wealth holdings as Cagetti (2003), here all percentiles of the wealth distribution from 10 to 90 are considered. Increasing the number of moments to match leads to estimates of the discount factor which are an order of magnitude more precise than in Cagetti (2003). The estimate for the discount factor, at $\hat{\delta} = 0.985$, is at the upper bound of the estimates in Cagetti (2003), while the estimated risk aversion parameter $\hat{\sigma} = 1.08$ is only a third to one sixth as large as Cagetti's estimate, depending on the subgroup under consideration.

Exercises like the ones by Cagetti (2003) and Hintermaier and Koeniger (2011) repeatedly come to one conclusion: while a simple life-cycle incomplete markets model with idiosyncratic income shocks calibrated from income data can match parts of the wealth distribution well, and generate the correct ordering of inequality in wealth, income, and consumption – wealth being more unequally distributed than income, which in turn is more unequally distributed than consumption – it fails to capture the extremely high dispersion in wealth, especially at the top of the distribution[6].

[6]Indeed, this is the reason cited in Hintermaier and Koeniger (2011) for excluding the top decile of the wealth distribution from the targeted moments: the model has no chance of capturing the extremely high net worth of the richest households, which exceeds 150 times average yearly income in the 2007 SCF.

A straightforward way of improving the fit of the more standard model is employing an income process that features large persistent shocks with low probability, as first popularised by Castaneda et al. (2003). Their specification of the income process is a four-state Markov chain, the highest state of which is only reached with very low probability, has a persistence of about five years, and implies an income 1000 times higher than median income. In this setup, it is evident that simple consumption smoothing considerations lead to very high savings rates for rich households, which in turn lead to a large wealth concentration at the top end of the distribution. It is however questionable to what extent models relying on this type of income process, which cannot be reconciled with the evidence from micro-level surveys on household incomes, can be used to inform policy analysis; a recent example of this problem can be found in the work of Kindermann and Krueger (2014), who investigate optimal labour income taxation in an Aiyagari-Bewley-Huggett style model featuring a similar income process and validating their model by fitting the top tail of the empirical wealth distribution. Unsurprisingly, they find very high optimal tax rates of around 90 percent on top earners, however this result is entirely driven by the income process used and subsequent work by Badel and Huggett (2014) demonstrates that optimal tax rates are significantly lower if one includes an earnings process based on human capital formation, which is parametrised to mirror the empirical evolution of earnings dispersion. This shows that a model that fits the wealth distribution is not necessarily suited for policy evaluation, especially if the good fit is the artefact of model assumptions that have little empirical support and gives reason to include more realistic features of the economic environment into the model which might help to explain observed patterns in the data. A number of researchers have extended the baseline model in various dimensions, some of which we will discuss here[7].

A more realistic version on the role of the household income process in shaping the wealth distribution is the inclusion of entrepreneurial activity as an alternative to labour income. Quadrini (2000) is an early attempt to include business income in the model. His economy features infinitely lived households, that have the opportunity to undertake entrepreneurial activity, but need to save up capital in order to start a business first. After having started the business, these agents face substantially higher risk than working households, a fact that combined with high borrowing costs and infinite lives leads to wealth accumulation at the top of the distribution as large as in the data. Cagetti and DeNardi (2006) improve on this model by allowing for an endogenous choice in the amount of capital invested in the business, and in their model the potentially high rates of return on business activity are the main factor affecting the right tail of the wealth distribution. This aspect makes their model a close cousins to models that feature different rates of returns for different asset classes, which we turn to later. Cagetti and DeNardi (2006) also provide an empirical rationale for the modelling of entrepreneurial activity, using SCF data to show that amongst the wealthiest 1% of households, 81% are business owners or self employed, although this group of households only accounts for 17% of all households. They also show that amongst business owners, mean and median wealth are higher for those not actively engaged in managing the business, providing support for models that feature an intergenerational transfer of assets, which we turn to next.

[7] The discussion here draws on the work by DeNardi (2015).

A successful line of research extends the model by moving from a simple life- cycle perspective to an overlapping generations (OLG) model, in which inequalities can be transmitted across generations and hence accumulate over time. This transmission mechanism can work through either assets directly, by adding a bequest motive to the agents utility function which prevents them from drawing down assets in old age, or through heritability of human capital in the form of skills or learning ability. The role of inheritance rose to prominence in the empirical literature with a dispute between Kotlikoff and Summers (1981), who estimate that around 80% of total wealth is inherited, while just 20% is the result of life-cycle saving, and Modigliani (1986), who argues that the role of the sources of wealth accumulations is exactly reversed. Recently, Thomas Piketty and a number of co-authors (Piketty 2011, Piketty et al. 2014, Piketty and Zucman 2015) revive this debate using long-run time series from France and drawing on other work from the UK and Germany, finding large variations in the annual flow of inheritance as a share of total wealth, but concluding that the overall importance of inheritance in shaping the wealth distribution is closer to Kotlikoff and Summer's estimates than to Modigliani's. The correct estimation of the role of inherited wealth is further complicated by the possibility of inter vivos transfers, which are not captured by inheritance tax data. This point is made forcefully by Gale and Scholz (1994), who use data on transfers from the SCF to estimate that 20 percent of aggregate wealth is passed on across generations via inter vivos transfers (compared to 31 percent as inheritances by their accounting). While the empirical estimation of intergenerational transfers of financial assets is not entirely straightforward, the question of the intergenerational transfer of ability is even more complicated. Researchers have adopted a wide range of specifications for modelling this transfer, based on models of parental investments in their childrens' education, or taken the short cut of directly assuming that children receive draws from productivity distributions, the mean and/or variance of which are directly linked to the parental realisation of productivity. DeNardi (2004) examines both bequests and intergenerational transmission of ability in tandem, and shows that while the model fit is vastly improved by this mechanism, it still misses the very high concentration of wealth in the top percentile of the wealth distribution.

Another line of work considers the role of preferences in driving inequality in wealth accumulation. The obvious way to affect the distribution of wealth through preferences is by letting the discount factor vary across agents, an idea that finds empirical support in work by Lawrance (1991), who finds significant heterogeneity in time preferences rates between poor and rich households using an Euler equation based regression approach on PSID income and consumption data. The first work to leverage differential discount factors to increase wealth dispersion in the Aiyagari-Bewley-Hugget framework is the seminal paper by Krusell et al. (1998), who experiment with three groups of agents exhibiting discount factors between 0.9858 and 0.993. Even with this seemingly small dispersion in preferences, the inequality in wealth holdings in the model rises dramatically, with the share of wealth held by the richest 1% of households increasing from three to 24 percent, and the Gini coefficient increasing from 0.25 to 0.82. This finding is corroborated in recent work by Carroll et al. (2014), who show that a model with a slightly higher dispersion in δ than in Krusell et al. (1998) can match both the Lorenz curve of net wealth and financial wealth almost exactly. Cozzi (2014) considers the implications of varying the other deep

parameter in the preference structure, risk aversion. He solves a model in which the population of agents has a mean risk aversion of 1.07, with a variance of 0.76, and shows that including this dimension of heterogeneity helps the model fit the data almost as well as the stochastic-delta model of Krusell and Smith, although it misses the concentration in the top percentile. Importantly, this model implies a significantly lower discount factor between 0.87 and 0.89 depending on the calibration. Interestingly, Cozzi combines his analysis with the estimation of income processes similar to the restricted income processes we will estimate in chapter 3 for subsamples of the PSID grouped by risk aversion, and finds significant heterogeneity in the persistence of the permanent shock to income, estimated at 0.947 for the less risk averse subgroup and 0.935 for the group with high risk aversion.
Going one step further than simply adjusting the parameters of the standard CRRA utility function, Diaz et al. (2003) depart from this utility function altogether and investigate the role of habits in the utility function, first introduced into the macroeconomic literature by Fuhrer (2000) in the context of a DSGE model of monetary policy. They show that while habits induce a significant increase in precautionary savings in the economy, they do not help to bring the model closer to the empirical dispersion of wealth holdings, and on the contrary lower the Gini coefficient compared to a model without habit formation.

As alluded to in the discussion of models with entrepreneurial activity, there is also a literature that increases wealth inequality predicted by Bewley style models by allowing for differential rates of return, a feature that finds support in a vast macro-finance literature on the equity premium puzzle (for a survey see Siegel and

Thaler (1997)), as well as the literature on households' portfolio choices (Mankiw and Zeldes (1991) discuss the equity market participation of US households, while Guiso et al. (2003) review the European evidence). Benhabib et al. (2011) devise a peculiar model of differing rates of return, where the difference don't arise across asset classes, but across generations, with each generation of a household drawing an idiosyncratic interest rate for its portfolio, which prevails for the entire span of its life[8]. Combined with altruism for future generations, this setup generates a consumption smoothing motive across generations, with generations of a household that draw a high rate of return accumulating assets to increase consumption of its descendants. Benhabib et al. also offer some empirical support for the relevance of heterogeneity in rates of return, citing a standard deviation of rates of return for housing equity of 14%, and an even higher standard deviation in the rates of return for business equity, to asset classes which account for 28.2 and 27 percent of total US household wealth, respectively.

Lastly, there might be institutional factors that exert differential influences on the savings behaviour of different agents. A prominent example of this can be found in the work of Hubbard et al. (1995), who show that in a model which includes a social security system based on asset-based means testing, something that can be found in virtually all advanced economies, there is a strong incentive for poor households not to accumulate any wealth, which increases wealth dispersion by lowering the wealth holdings at the bottom of the distribution. Other government programs such as Medicaid in the US might play a role in shaping the dissaving behaviour of elderly households, which the standard model also

[8]Highlighting the similarities between models of entrepreneurial activity and those featuring different rates of return, Benhabib et al. motivate the inclusion of stochastic rates of return as an attempt to capture entrepreneurial risk.

has problems in replicating (see DeNardi et al. (2009) and 2010, DeNardi et al. (2015)).

2.9 Conclusion

This chapter provides an overview of stylised facts about the wealth distributions in a number of advanced economies and presents various approaches to build economic models which can account for these stylised facts. It became clear that while saving for retirement is the main driver of wealth accumulation for large parts of the population, other factors need to be taken into consideration to explain the tails of the distribution and the behaviour of young households. Crucial aspects of an economic model of the wealth distribution are the risks households are facing – both on the income and the expenditure side – and the financial markets available to them to insure themselves against those risks and earn returns on their savings. Finally, the far right tail of the wealth distribution seems to be driven by factors beyond this, with modelling attempts featuring a vast array of different ingredients succeeding in matching the distribution of wealth even for the richest house- holds. Given that vastly different approaches manage to fit the distribution, it is fair to say that so far there is no consensus on which mechanism is the most important one to include, and that all attempts to match the observed dispersion of wealth based on one of those mechanisms likely overstate the contribution of that particular mechanism, as so far no attempts at building an overlapping generations stochastic-beta model featuring differential rates of return, a realistic tax and benefit system, entrepreneurial activity, intergenerational transmission of financial wealth and ability, and human capital formation has been made. Some progress is being made in this direction, e.g. in DeNardi and Yang (2015), who combine intergenerational transfers of wealth and ability with an income process exhibiting higher income risk for rich households. Furthermore, virtually all of the papers discussed in this chapter rely on a variation of a simple AR(1) income process, ignoring recent evidence on income processes from large administrative data sets, to be discussed in chapter 3. This means that the implications of heterogeneous income processes for the wealth distribution are not well understood, a gap in the literature that 4 will attempt to address.

Chapter 3

The Effects of Profile Heterogeneity on Estimates of Income Risk

3.1 Introduction

As has become clear from the discussion in the preceding chapter, a crucial ingredient to any model of household savings is an estimate of the risk that households are facing in the form of their income process. Traditionally, researchers have relied on a parsimonious AR(1) specification with a transitory and a persistent shock component, which can be represented as a Markov chain and thus helps to ease the computational burden. Recent research has cast doubt on the ability of this specification to accurately capture the risk faced by households in the labour market though, and advances in computational capabilities have allowed to solve models with larger state spaces, so that there is a renewed interest in estimating richer statistical processes for household income.

The labour economics literature of income processes has long attempted to model household earnings dynamics using a variety of rich time series models with different AR and MA specifications. Early attempts to exploit longitudinal data on household's income include the seminal work of MaCurdy (1982), who fits ARMA processes to the income levels of a sample of prime age males from the first ten waves of the PSID and concludes that the data is best described by either an ARMA(1,2) or an ARMA(2,1) process; Abowd and Card (1987), who analyse data from the PSID, the NLS and SIME/DIME and settle for an MA(2) description of the data as most appropriate. Both MaCurdy (1982) and Abowd and Card (1987) conclude that the autoregressive component of the stochastic process describing income residuals has to have a unit root, a conclusion that is called into question by Baker (1997), who develops econometric tests that reject a specification with $\rho = 1$, and favour a specification with what he calls heterogeneous profiles, that is, an individual specific slope component in the income process. This approach had been previously applied in longitudinal data on American scientists[9] by Weiss and Lillard (1978) and in data on 279 Swedish scientists by Hause (1980). While these papers rely on a deterministic structure for individual wage growth over the life-cycle, Guvenen (2009) offers a model that fuses these approaches, including both deterministic components for the level and slope of income, as well as a stochastic AR(1) component delivering persistent shocks. As we will base our analysis on this model, we defer the detailed model description to the next section.

While most of the work discussed so far has relied on the use of survey data of income, which is plagued by measurement error and hence cannot correctly

[9]The National Science Foundation's Register of Technical and Scientific Personnel, a dataset comprised of bi-yearly income observations on Ph.D. holders in the STEM fields.

identify the variance of transitory income shocks, in recent years researchers have been able to make use of the huge data base of the US Social Security administration, which offers exact data on incomes of millions of American workers over long periods of time. The first papers to make use of this data were Kopczuk et al. (2010), who focus on the evolution of cross-sectional income inequality over time and the distinction between permanent and transitory shocks to income, and DeBacker et al. (2013), who use a similar data set of tax returns to answer a very similar question – we will return to the implications of their findings for macroeconomic models in chapter 3. More interesting in the present context is a recent paper by Guvenen et al. (2015), who use the Master Earnings File of the Social Security administration for the years 1978 to 2010 to construct an extremely large panel of income observations for a sample of 10% of all US workers that were issued a Social Security number. From this data set, the authors conclude that the distribution of income shocks is not normal, with a kurtosis ten- to fifteen times that of a Normal distribution. Fitting processes similar to that in Guvenen (2009) to the data, they conclude that the data is best described by a model including heterogeneity in individual specific growth rates and a mixture of (at least) two independent AR(1) processes with different innovation variance. Some more recent papers take the opposite stance though and argue that profile heterogeneity is in fact not present in the variance-covariance structure of income data. Hoffmann (2013) uses administrative records from the German Institut für Arbeitsmarkt- und Berufsforschung (IAB), which allows to construct individual-specific earnings histories for up to 120 quarters and is fairly large, representing a 2% sample of all German salaried employees. Given the structure of the data, it is possible to control for age and cohort effects better than in the

PSID, were small sample sizes force aggregation of age groups. Hoffman finds intercept heterogeneity (σ_α^2) to be an important feature when trying to fit the data irrespective across all specification of income processes under consideration, but argues that heterogeneity in income growth rates becomes insignificant once the variance of the initial value of the persistent component is adequately controlled for. Along similar lines, Hryshko (2012) conducts Monte Carlo simulations to show that if a misspecified HIP model is estimated on a synthetic dataset generated from an underlying process with $\sigma_\beta^2 = 0$, an econometrician will generally find statistically significant levels of profile heterogeneity. Finally, some authors have attempted to extend the basic ARMA model in other directions, adding e.g. ARCH effects to capture stochastic volatility in income innovations (Meghir and Pistaferri 2004), allowing for individual-specific income *processes*, rather than simply different means and variances for the same process (Browning et al. 2010) It is thus fair to say that the literature has not yet reached a firm conclusion on the correct specification of the income process households are facing. This chapter will undertake a modest attempt at adding to the evidence by estimating RIP and HIP processes on different samples of income data. To our knowledge, this is the first study to use all available waves for the PSID, ranging from 1968 to 2013, and the first study to estimate HIP processes from data coming from the British Household Panel Study (BHPS).

3.2 The statistical model

To inform the simulations in the following chapter, this thesis will rely on an estimated heterogeneous income profiles (HIP) income process in the spirit of

Guvenen (2009). The process to be estimated is of the form

$$y^i_{h,t} = g(\theta_t, \boldsymbol{X}^i_{h,t}) + \alpha^i + \beta^i h + z^i_{h,t} + \phi\varepsilon^i_{h,t} \tag{3.1}$$

$$z_{h,t} = \rho z_{h-1,t-1} + \pi_t \eta^i_{h,t} \tag{3.2}$$

where $y^i_{h,t}$ are the log earnings of individual i, who has h years of labour market experience in period t. The function $g()$ is assumed to be a cubic polynomial in experience, while the individual specific parameters α^i and β^i – modelled as random variables with mean zero and variance σ^2_α and σ^2_β, respectively – capture the cross-sectional profile heterogeneity. $z^i_{h,t}$ is an AR(1) process with persistence ρ and innovation variance $\eta^i_{h,t}$, which captures persistent shocks to income, while $\varepsilon^i_{h,t}$ is a purely transitory shock. Both η^i and ε^i are mean-zero i.i.d random variables with variances σ^2_η and σ^2_ε, respectively. As discussed above, the variances of both permanent and transitory shocks have seen large swings over the past decades, to capture this we are allowing for time-variation in the innovation variance (denoted π_t for the innovation to the persistent shock component and φ_t for the transitory counterpart).

To estimate the parameters of the model, an equally weighted minimum distance estimator is used to minimise the distance between the empirically observed variance-covariance structure of residual earnings (defined as $\tilde{y} \equiv y^i_{h,t} - g(\theta_t, \boldsymbol{X}^i_{h,t})$ and the variance-covariance structure implied by the model. In the present context, this strategy has first been employed by Baker (1997), who estimates a very similar model to the one described above, although the approach has been used before for estimating other models in labour economics, e.g. Abowd and

Table 3.1: Previous estimates of profile heterogeneity in different data sets

	ρ	σ_η^2	σ_ε^2	σ_α^2	σ_β^2	$cov(\alpha,\beta)$
Guvenen (2009)	0.821	0.029	0.047	0.022	0.00038	-0.23
Baker (1997)	0.423	0.089	–	0.355	0.00081	-0.014
Haider (2001)	0.639	0.057–0.166	–	0.295	0.00041	-0.0083

Card (1987). Table 3.1 summarizes findings of earlier papers. Our model implies theoretical variances and covariances given by:

$$\mathrm{Var}(\tilde{y}^i_{h,t}) = \underbrace{\sigma_\alpha^2 + 2\sigma_{\alpha\beta}h + \sigma_\beta^2}_{\text{contribution of profile heterogeneity}} + \mathrm{Var}(z^i_{h,t}) + \phi_t^2\sigma_\varepsilon^2 \tag{3.3}$$

$$\mathrm{Cov}(\tilde{y}^i_{h,t}, \tilde{y}^i_{h+n,t+n}) = \underbrace{\sigma_\alpha^2 + 2\sigma_{\alpha\beta}(h+n) + \sigma_\beta^2} + \mathrm{Var}(z^i_{h,t}) + \phi_t^2\sigma_\varepsilon^2 \tag{3.4}$$

The empirical variance-covariance matrix underlying the estimation will be obtained by first calculating the covariance of residuals for each age-group in a given year, and then averaging over all age groups present in a given year. The theoretical counterpart is obtained by simply calculating the corresponding variances and covariances from the formulas above, and forming weighted averages over h with weights corresponding to the relative frequency of age-groups in the empirical data.

[9] All codes used in this chapter can be found on my GitHub page: psidJulia for the code used to merge all waves of the PSID, extract the variance-covariance matrix of residuals and fit the process using a minimum distance estimator, BHPStools for code that merges the 18 waves of the BHPS and creates the residual variance-covariance matrix.

3.3 Data

As we are interested in the variability of our estimates, we are estimating the process described both on PSID and BHPS data. PSID data has the advantage of providing a very long horizon (37 waves of data covering a total of 45 years), which allows for the analysis of sub-periods to examine changes over time. The BHPS, while more limited in time (18 waves of data covering 18 years) serves as a useful comparison, while also providing excellent measures of different measures of household incomes pre- and post taxes and transfers, which we will describe in more detail below. The data is taken from all available waves of the PSID, that is years 1968 to 2013 inclusive[10]. For our baseline estimation, to ease comparisons, we stick to the sample selection criteria used in Guvenen (2009), namely:

- Household heads between the ages of 20 and 64 inclusive
- Hourly labour earnings between $2 and $400 in 1993 prices
- Hours worked between 520 and 5110

For inclusion in our sample, an individual has to fulfil all of the above conditions for at least 20, not necessarily consecutive, years. These sample selection criteria leave us with 1685 individuals in our final sample [11]. The main variable of interest in the analysis is labour income, for which we use the series of variables starting with `V74` in 1968[12]. Hourly earnings are taken from the variable starting with `V337` in 1968, while hours worked are taken from the variable starting with `V47`.

[10]Note that the PSID income variable refers to income in the previous year, so when we talk about data from, e.g., year 1968, it is implied that we are referring to income in 1967.

[11]To create the longitudinal data set from the PSID cross-sections, we use the excellent *PSIDtools* package (Kohler, 2015)

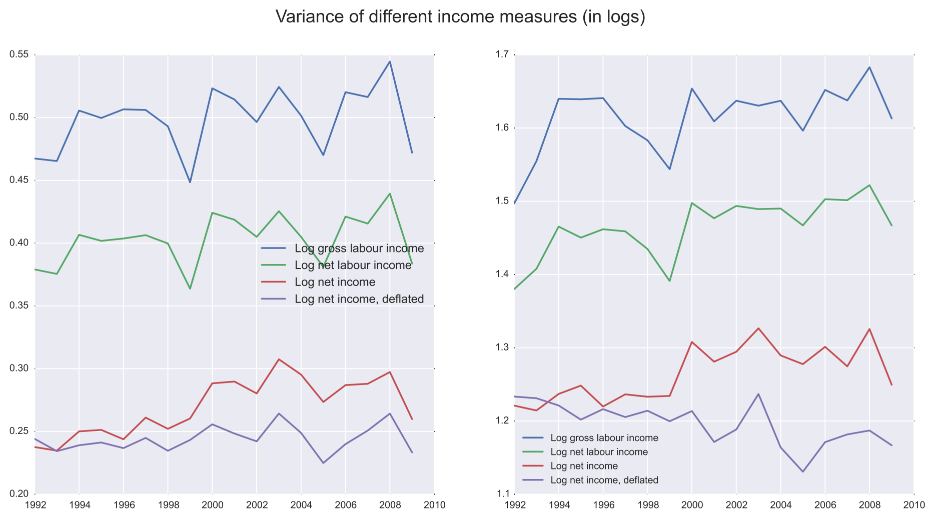
[12]A complete list of all variables used is available on my GitHub page.

To extract the deterministic life-cycle component of the is modelled as a cubic polynomial in experience, $g(\theta_t, \boldsymbol{X}^i_{h,t}) = \gamma_0 + \gamma_1 h + \gamma_2 h^2 + \gamma_3 h^3$. Labour market experience itself is constructed as potential experience from information on years of schooling.

The BHPS data we are using comes from all available waves, covering the time period from 1992 to 2008[13]. The raw data is then extended by the derived current annual and net household variable data set provided by Horacio Levy and Stephen Jenkins, described in Jenkins (2010). This data set includes information on household income that takes into account various government taxes and transfers, both at the individual and the household level. For our purposes, we will use gross labour income of the household, which is available in the original BHPS data set; net household labour income, which considers taxes and tax credits, national insurance contributions, and occupational pension contributions; and net household income, which adds investment income, pension income, and transfer income to net labour income. These three variables can be seen to represent different levels of insurance available to the household: as taxes and (up to a points) National Insurance contributions in Britain are progressive, they reduce the variability of the labour income process facing the household, while the benefits system, which includes housing benefit, job seekers benefits, disability insurance and various other payments, partially insures household income against unemployment and other catastrophic shocks. It is therefore expected that these measures of income imply less risk for the household than gross labour income, an effect that we will try to quantify below. As the time dimension is notably shorter than in the PSID, we only require households to be in the sample for five years, and

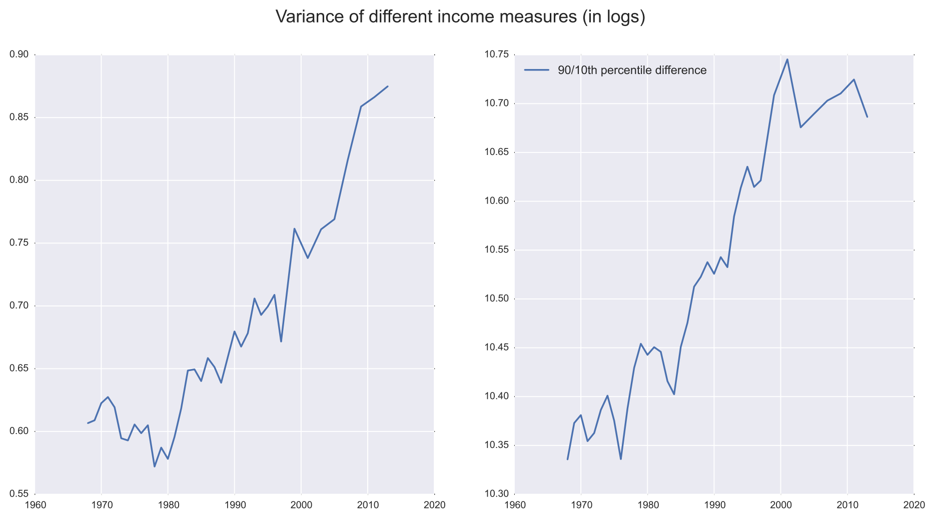
[13]To merge the BHPS data across waves, we use code provided by Vandendriessche (2015).

Figure 3.1: Variance of log income and 90/10 percentile width for our sample of BHPS households



consider up to ten lags for the covariances of residuals. While previous authors have highlighted the importance of higher order covariances for identification of HIP processes, our sample sizes unfortunately are so small that for some cohorts there are less than 10 observations at lags larger than five, so that considering more lags is impossible. As for the PSID, we obtain income residuals by regressing each measure of income on a cubic polynomial in experience, which is constructed from the school leaving age (or further education leaving age, where applicable). Figure 3.1 and 3.2 show trends in the variance and the inter-decile range, two widely used measures of income dispersion, for our selected sample of households. Both datasets exhibit considerable variation in the dispersion of income over the period under consideration, which motivates us to include time-varying variances for transitory and permanent shocks in the estimation.

Figure 3.2: Variance of log income and 90/10 percentile width for our sample of BHPS households

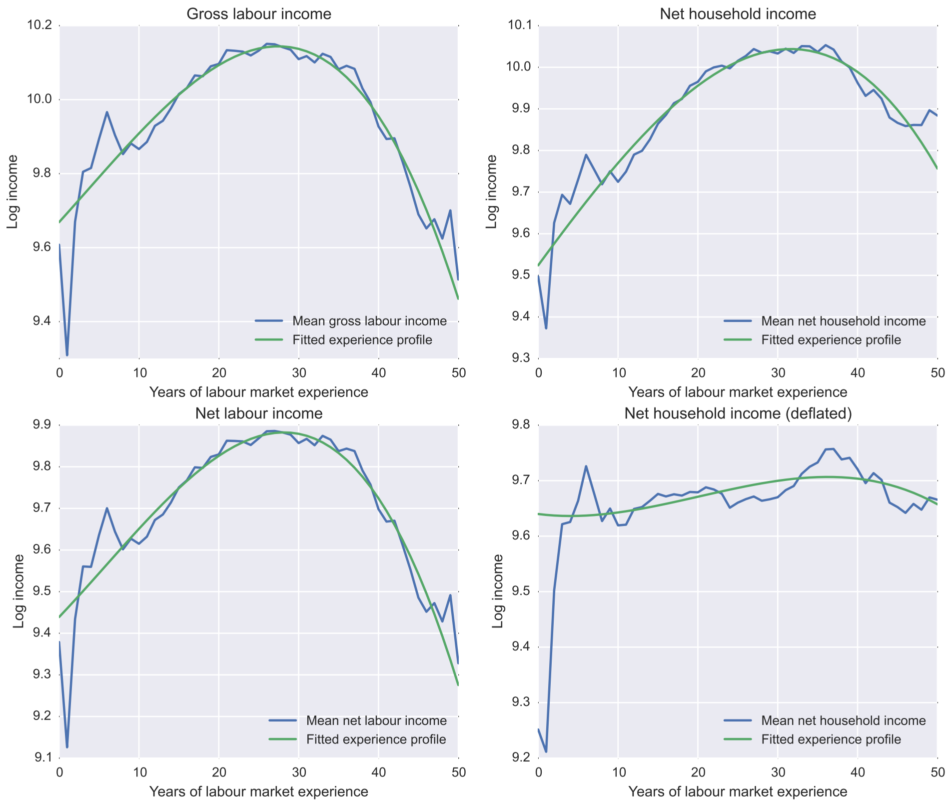


3.4 Results

Table 3.3 displays the results for the PSID sample of households. While the results can be considered to be qualitatively similar to those found in Guvenen (2009), interestingly the estimated dispersion of individual- specific growth rates declines from 0.00031 to 0.00025, a result that confirms the same finding in Hryshko (2012). Furthermore, the difference in estimated persistence for the HIP and RIP is much smaller than that reported by Guvenen, and closer to the values found in Hryshko (2012)[14]. Table 3.2 shows the results of estimating both RIP and HIP processes on our sample of households from the BHPS, using the four different income measures described previously. The results are largely

[14]To ensure broad correctness of our estimation procedure, we also downloaded and ran the code of Guvenen (2009) along with the original dataset used therein from the journal website. Unfortunately, we were unable to replicate the results reported in the paper with this code; the results turn out to be much closer to our results reported here. The author did not respond to repeated emails asking for clarification.

Figure 3.3: Log mean income and fitted experience profiles for the BHPS 1992 – 2009



unsurprising qualitatively, with the variance of persistent shocks declining from 0.1 for the most volatile process (gross labour earnings) to 0.07 for net labour earnings, to 0.027 for net household income. A similar pattern can be observed for the transitory shock, declining from 0.08 to 0.07 and 0.056, respectively. The estimates for the main parameter of interest, the cross-sectional dispersion in individual specific growth rates β^i behaves accordingly, dropping from 0.00032 (consistent with the findings for the main PSID sample in Guvenen (2009)), to 0.00019 and 0.00011. Interestingly, some of the decrease in the parameters that increase cross-sectional dispersion over the life-cycle of a cohort is offset by a rise in the cross-sectional inequality in intercepts, Var_α rises from 0.036 in the gross labour income sample to 0.052 in the net household income sample. The persistence of The most peculiar set of estimates obtains for the deflated and equivalized measure of net household income. Here, the RIP process shows a much lower persistence as would be expected, while the variance of persistence shocks is surprisingly higher than in then in the raw measure of net household income. The results for the HIP process indicate that the minimization routine hit the boundaries on both σ_β^2[15] and $\text{Cov}(\alpha, \beta)$. A possible explanation for the large difference in results compared to all other income processes can be seen in figure 3.3: as the equivalization largely removes the hump-shape of the experience profile in the data, the fitted regression line misses the sharp increase in income at the earliest stage of the life cycle, and instead takes a flat shape over the entire range. This necessarily implies an entirely different structure of residuals, with extremely large predicted residuals for the first years of working life. Given this,

[15]Which actually has a lower bound of 1e-6, so is not exactly zero.

we will not consider the estimates for this process in the rest of this thesis[16].

3.5 Discussion

Our estimates point to substantial uncertainty over the correct HIP process, adding to a literature that has found vastly different estimates for all of the main parameters. Further, we have documented that even applying the same estimation procedure to different subsamples of the same survey can deliver results that differ markedly. Lastly, our estimates based on different income measures from the BHPS underscore the importance of partial insurance when trying to estimate household income risk from the data. In the next chapter, we will explore the quantitative implications of these differences for wealth accumulation in life-cycle models with incomplete markets.

[16]For the case of the HIP process this isn't actually a choice, as the estimated parameters for σ^2_β and Covα, β imply a negative-definite variance covariance matrix for the bivariate Normal distribution from which α and β are drawn, making it impossible to simulate an income distribution using these estimates.

Table 3.2: Results for the BHPS sample 1992–2008, different measures of household income

	ρ	σ^2_η	σ^2_ε	σ^2_α	σ^2_β	$\sigma_{\alpha\beta}$
Restricted income process: $\sigma^2_\beta \overset{!}{=} 0$						
Gross labour income	0.925	0.045	0.135	0.0	–	–
Net labour income	0.867	0.065	0.077	0.017	–	–
Net household income	0.921	0.026	0.046	0.012	–	–
Net household income (deflated)	0.817	0.038	0.038	0.084	–	–
Heterogeneous income process; σ^2_β unrestricted						
Gross labour income	0.719	0.106	0.080	0.036	0.00032	-0.51
Net labour income	0.808	0.073	0.070	0.032	0.00019	-0.59
Net household income	0.857	0.027	0.056	0.052	0.00011	-0.42
Net household income (deflated)	0.812	0.039	0.042	0.100	0.0	-1.0

Table 3.3: Results for the PSID, different sample periods.

	ρ	σ^2_η	σ^2_ε	σ^2_α	σ^2_β	$\sigma_{\alpha\beta}$
Restricted income process: $\sigma^2_\beta \overset{!}{=} 0$						
1968-1996 sample	0.932	0.010	0.036	0.084	–	–
1968-2013 sample	0.920	0.014	0.067	0.076	–	–
1968-1986 sample	0.960	0.015	0.061	0.058	–	–
1987-2013 sample	0.939	0.017	0.095	0.110	–	–
Heterogeneous income process: σ^2_β unrestricted						
1968-1996 sample	0.853	0.013	0.030	0.030	0.00031	–0.30
1968-2013 sample	0.839	0.017	0.064	0.047	0.00026	–0.32
1968-1986 sample	0.885	0.013	0.043	0.110	0.00028	-0.42
1987-2013 sample	0.854	0.032	0.085	0.097	0.00025	-0.31

Chapter 4

Wealth Distributions in Heterogeneous Income Process Models with Learning

4.1 Introduction

The preceding two chapters have pointed out the importance of developing quantitative economic models of household saving and the important interconnections between savings, insurance, and income risk. This chapter builds on the discussion of the previous chapter by calibrating a model of household saving featuring a heterogeneous income process to empirical data on the wealth distribution and assessing the models fit under different parametrisations of the income process. The main finding is that including profile heterogeneity significantly worsens the models ability to fit the data moments, a finding which is robust across all income processes estimated in the previous chapter. We then conduct comparative statics exercises to explore what drives this result. These exercises show that it is precisely the main difference that separates HIP from RIP processes – the lower estimated persistence and variance of transitory shocks – which keeps the model from matching the shape of the empirical wealth distribution. While lifetime inequality in an income distribution simulated from an HIP process with modest amounts of profile heterogeneity is just as high as that found in an income distribution simulated from a persistent AR(1) process with high innovation variance, the different mechanism generating this inequality crucially alters household savings behaviour.

The model employed in this chapter is a standard incomplete markets life-cycle model of household consumption and savings as described in chapter 2, with the addition of a learning mechanism for household income, as first used by Guvenen (2007). While this model has recently been used to study household's portfolio choices (Chang et al. 2013) and the joint evolution of income and consumption inequality in a rich dynamic model featuring informal insurance mechanisms (Guvenen and Smith 2014), the implications of heterogeneous income processes and learning for the aggregate wealth distribution have so far not been examined to the best of our knowledge. Although a very early working paper version of Guvenen (2007) briefly comments on the wealth distribution that obtains in a model of profile heterogeneity and learning[17], noting that overall wealth accumulation is significantly higher than in a model without profile uncertainty, to the best of our knowledge no further investigations into the ability of life-cycle models with profile heterogeneity to match the wealth distribution have been

[17]At the time of writing, the working paper can be accessed at `http://www.usc.edu/schools/business/FBE/seminars/papers/M_5-18-04_GUVENEN-Labrisk04.pdf`.

undertaken. Given the recent evidence discussed in chapter 3 pointing towards profile heterogeneity as an important feature of real life income processes, it is important to understand what including these income processes in life-cycle models – arguably making them reflect better the actual risk facing households in reality – means for the wealth distribution predicted by the model. While chapter 2 showed that there are many promising extensions to the standard model making it conform better to the empirical wealth distribution, virtually all papers discussed there rely on a simple AR(1) process with persistent and transitory shocks for the modelling of income risk, a process that, if the HIP specification estimated in chapter 3 is correct, significantly overstates both the magnitude and the persistence of permanent shocks to household income. Therefore, this chapter examines theoretical wealth distributions coming out of models relying on the HIP process for household income, under the assumption that households are imperfectly informed about the parameters of their individual income process, but able to learn them over the course of their working life. Using the parameter estimates from chapter 3, it explores the implications of changes over time and across household income measures for the predicted wealth distribution. Following Hintermaier and Koeniger (2011), it then uses a minimum distance estimator to calibrate discount factor and risk aversion in order to minimise the difference between wealth holdings at different percentiles of the wealth distribution in the model and the data. Then, comparative statics exercises are performed to isolate the role of the different parameters of the HIP process in determining the shape of the income distribution. Finally, the sensitivity of the model to changes in agents' initial beliefs and the role of systematic mistakes in beliefs is investigated.

4.2 The Model

Consumers maximize

$$E_0\left[\sum_{t=0}^{T}\beta^t\frac{c_t^{1-\sigma}}{1-\sigma}\right] \tag{4.1}$$

s.t.

$$a_{t+1} = (1+r)a_t + y_t - c_t \tag{4.2}$$

$$y_t^i = g(\theta^0, X_t^i) + f(\theta^i, X_t^i) + z_t^i + \epsilon_t^i \tag{4.3}$$

$$a_{t+1} \geq \underline{a} \tag{4.4}$$

where c_t is consumption in period t, a_t are asset holdings subject to a borrowing constraint $\underline{a}$, and y_t^i is individual income, which follows the heterogeneous income specification in logs discussed in chapter 3:

$$y_t^i = g(\theta^0, X_t^i) + f(\theta^i, X_t^i) + z_t^i + \epsilon_t^i$$

Here, $g(\theta^0, X_t^i)$ captures age effects and individual specific characteristics such as education, z_t^i is an autoregressive process of order one, and $f(\cdot)$ is an individual specific function that plays the decisive role in introducing heterogeneity and learning in the model.

$$\begin{aligned} f(\theta^i, X_t^i) &= \alpha^i + \beta^i t \\ z_t^i &= \rho z_{t-1}^i + \eta_t^i \\ \theta^i &\sim N\left[\begin{pmatrix} 0 \\ \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \\ \sigma_{\alpha\beta} & \sigma_\beta^2 \end{pmatrix}\right] \end{aligned}$$

The parameters α and β are randomly distributed over the population and govern the evolution of lifetime income over time. Furthermore, they are unknown to individuals upon entering the labour market, meaning that in order to calculate an expected lifetime income to base consumption choices on, consumers in the model have to form beliefs over the values of their individual parameters. Here again we follow Guvenen in assuming that these beliefs are formed optimally in a Bayesian fashion, which means solving a Kalman filtering problem. Denoting by S_{t+1}^i the vector of parameters α^i, β^i and z_{t+1}^i and by F the coefficient vector in the state space representation, the evolution is governed by the law of motion:

$$\hat{S}_{t|t}^i = \hat{S}_{t|t-1}^i + P_{t|t-1} H_t [H_t' P_{t|t-1} H_t + R]^{-1} (y_t^i - H_t' \hat{S}_{t|t-1}^i) \tag{4.5}$$

$$\hat{S}_{t+1|t}^i = F \hat{S}_{t|t}^i \tag{4.6}$$

where we denote by $\hat{S}_{t|t}^i$ the optimal belief about the individual specific parameters of the income process in period t after having observed the realisation of y_t^i, and by $\hat{S}_{t+1|t}^i$ the optimal forecast based on those beliefs, assuming that the transition matrix F is known to the household. $P_{t|t}$ is the variance-covariance matrix of $\hat{S}_{t|t}^i$

and R is the variance of the transitory shock. A similar expression can be derived for the evolution of $P_{t+1|t}$:

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H_t[H_t'P_{t|t-1}H_t + R]^{-1}H_t'P_{t|t-1} \tag{4.7}$$

$$P_{t+1|t} = FP_{t|t}F' + Q \tag{4.8}$$

With Q denoting the covariance matrix of the innovation in the state space representation of $\hat{S}^i_{t+1|t}$ (which is basically the innovation in the AR(1) component of earnings). Given this formulation for the evolution of beliefs, we can write the recursive version of our maximization problem as:

$$V_t(x_t, \hat{S}^i_{t|t}) = \max_{\{c^i_t, a^i_{t+1}\}} \left\{ u(c_t) + \mathbb{E}_t \left[V_{t+1}(x_{t+1}, \hat{S}^i_{t+1|t+1}|\hat{S}^i_{t|t}) \right] \right\} \tag{4.9}$$

which again has to be solved subject to the constraints, 4.2, 4.4 and 4.5-4.8. Note that given this formulation of the problem, all state variables appearing in the continuation value function on the right- hand side of the Bellman equation are functions of the realisation of income next period, so that the expectation in 4.9 has to be taken only with respect to $\hat{y}^i_{t+1}$. The distribution of next period's expectation of income is known exactly, conditional on current beliefs:

$$\hat{y}^i_{t+1} \sim \mathcal{N}(\hat{\alpha}_{t|t} + (t+1)\hat{\beta}^i_{t|t} + \rho\hat{z}_{t|t}, \sigma^2_\alpha + t^2\sigma^2_\beta + 2t\sigma_{\alpha\beta} + \sigma^2_\eta + \sigma^2_\varepsilon$$

An important issue when trying to match empirical wealth distributions is the specification of the pension system. The household problem during retirement is straightforward to solve in the absence of uncertainty; it is given by

$$V_t^R(a_t, \bar{y}) = \max_{c_t, a_{t+1}} u(c_t) + \delta V_{t+1}^R(a_{t+1}, y^R) \tag{4.10}$$

s.t.

$$a_{t+1} = (1+r)a_t + \bar{y} - c_t \tag{4.11}$$

$$y^R = M(\bar{y}, \bar{Y}) \tag{4.12}$$

$$a_{t+1} \geq \underline{a} \tag{4.13}$$

where M is a benefit function that emulates the US Social Security system and is specified following much of the literature on life-cycle models (cp. Storesletten et al. (2004), Hintermaier and Koeniger (2011), Guvenen and Smith (2014), amongst others) as a function depending on average lifetime income of an individual, $\bar{y}$, relative to the economy-wide average income $\bar{Y}$:

$$y^P = \begin{cases} 0.9\bar{y} & \text{if} \bar{y} < 0.3\bar{Y} \\ 0.27 + 0.32(\bar{y} - 0.3) & \text{if} \bar{y} \leq 2.0\bar{Y} \\ 0.814 + 0.15(\bar{y} - 2.0) & \text{if} \bar{y} \leq 4.1\bar{Y} \\ 1.129\bar{Y} & \text{if} \bar{y} > 4.1\bar{Y} \end{cases}$$

Note that this system attenuates the inequality in lifetime income created by the stochastic process for income by providing higher replacement rates for poor households than for rich households. To avoid adding another state variable to the model, we replace the true value of $\bar{y}$ by an estimate derived from running the cross-sectional regression

4.2.1 Computational Algorithm

To solve the model, we adopt a strategy similar to that in Guvenen and Smith (2014). After drawing an income distribution and simulating agent's learning given a set of initial beliefs, we construct a three-point grid for $\hat{\alpha}$, a fifteen-point grid for $\hat{\beta}$ and a seven-point grid for $\hat{z}$, all linearly spaced ranging from the lowest to the highest belief coming out of the simulation of agent's learning process. For wealth, we choose 40 grid points, exponentially spaced with a higher concentration of points at low levels of wealth. The household's pension problem can be solved analytically, while the household's working life problem is solved recursively on all grid points in the four- dimensional state space. To evaluate the continuation value function on the right- hand side of the Bellman equation, we employ quadrilinear interpolation combined with Gauss-Hermite quadrature on ten nodes for the numerical integration [18]. In the simulation step, we initialise household wealth holdings by drawing from the empirical wealth distribution for 23 to 25 year old households from the Survey of Consumer Finances, data that is available in Hintermaier and Koeniger (2011). We check the sensitivity of model results to this choice by comparing them to the alternative of zero wealth holdings at age 20 for all households and confirm both that there is no qualitative difference in model results, and that the quantitative differences are negligible. All codes used in this chapter can be found in my GitHub repository LearningModels.

[18] In particular, the linear interpolation was performed using the `ApproXD.jl` package (Oswald, 2014), while the Gauss-Hermite nodes were derived using the `FastGaussQuadrature.jl` package (Townsend, 2015).

4.3 Quantitative Results

To take the model to the data, we follow the strategy in Hintermaier and Koeniger (2011) and calibrate the model using a minimum distance estimator that minimizes the difference between wealth holdings at percentiles 10 to 90 of the net wealth distribution for different ages. The values for the SCF can be readily obtained from the code of Hintermaier and Koeniger (2011), while we use the UK Wealth and Asset survey to derive similar statistics for the UK for fitting the model when the income process is derived from BHPS data. The target moments in the data are the wealth holdings at percentiles 10 to 90 for prime age households (ages 26 to 55), as well as for young (ages 26 to 35), middle aged (ages 36 to 45) and old (46 to 55) households[19], which gives us a total of 324 moments to match. Stacking all of these moments into a vector μ, and denoting the vector of percentile wealth holdings for households simulated from the model by θ, the minimum distance estimator is

$$\min_{\delta,\sigma} \theta' I \theta$$

Hintermaier and Koeniger (2011) derive a normality result for the estimates obtained from this estimator, which allows us to compute standard errors as the main diagonal of the inverse of the squared Jacobian, $trace((J'J)^{(} - 1)$, where

$$J = \left[\frac{\partial \theta}{\partial \delta} \frac{\partial \theta}{\partial \sigma} \right]$$

As our baseline estimate, we fit the HIP process estimated from the full sample of PSID households from 1968 to 2013 in chapter 3. From the minimum distance

[19] When using WAS data, the age categories are all shifted back by one year (that is, young households are aged 25 to 34), as this is the categorisation used in the survey.

estimation, we obtain a discount factor of $\delta =$ and a risk aversion parameter $\sigma =$ as the best fit, although the value of the objective function at the minimum is still very high. The reason for this can be seen graphically in figure 4.1 and figure 4.2: the model entirely misses the shape of the empirical wealth distribution, predicting too little wealth at the low end of the distribution, too high wealth accumulation for households between the 20th and 70th percentile, and too small wealth holdings again for households at the highest percentiles [20]. Breaking the result down by age groups, we see that young households are much poorer in the model than they are in the data, while older cohorts display much higher savings.

From these results it is obvious that the shape of the wealth distribution predicted by the model is so fundamentally wrong that none of the processes estimated in chapter 3 will stand a chance of matching the empirical wealth distribution. In the interest of brevity, we therefore skip a graphical presentation of the model fit for the processes estimated from BHPS data to the empirical wealth distribution in the WAS, and only report the results of the minimum distance estimation in table 4.1. It can be seen that all estimates of the discount factor are significantly lower than those obtained by Hintermaier and Koeniger (2011), an artefact of the fact that overall wealth accumulation in the life-cycle model featuring an HIP income process is much higher than in a comparable model relying on a simple AR(1) income specification, as reported in the earlier working paper version of Guvenen (2007). In the first row, we also report the results for fitting a model using the HIP parameters estimated from the PSID data, but excluding the deterministic life cycle profile of earnings, $g_t(\theta_0, X_t^i)$, which is the

[20]Note that in all plots, we only display wealth holdings in model and data between the first and 90th percentile.

Figure 4.1: Baseline model (PSID HIP income process), calibrated fit

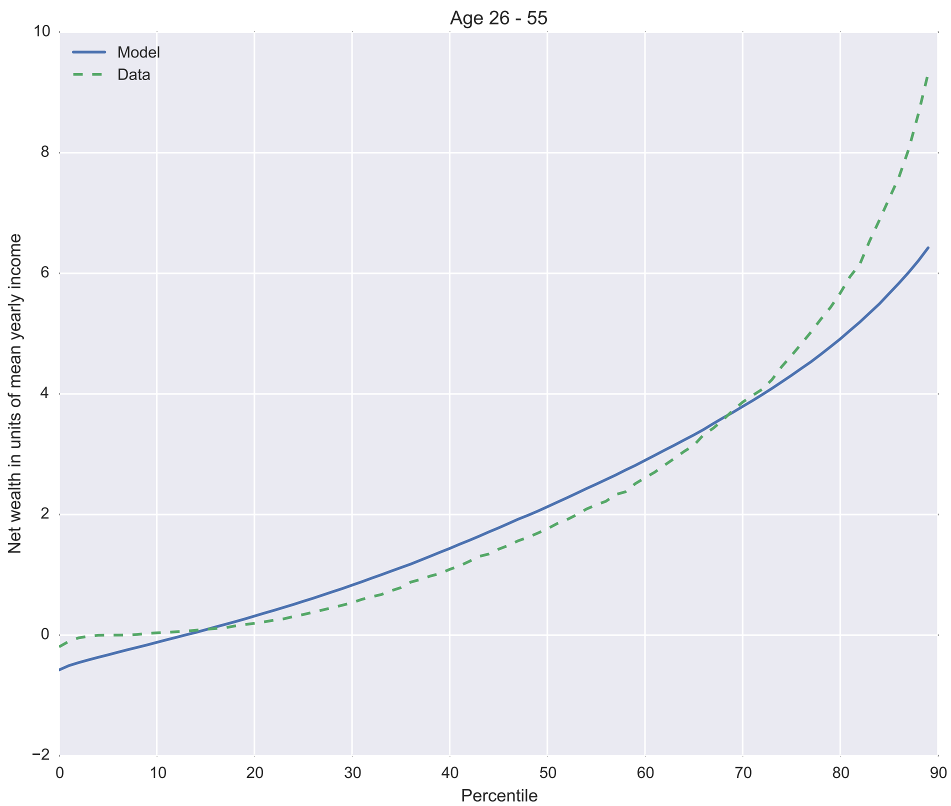
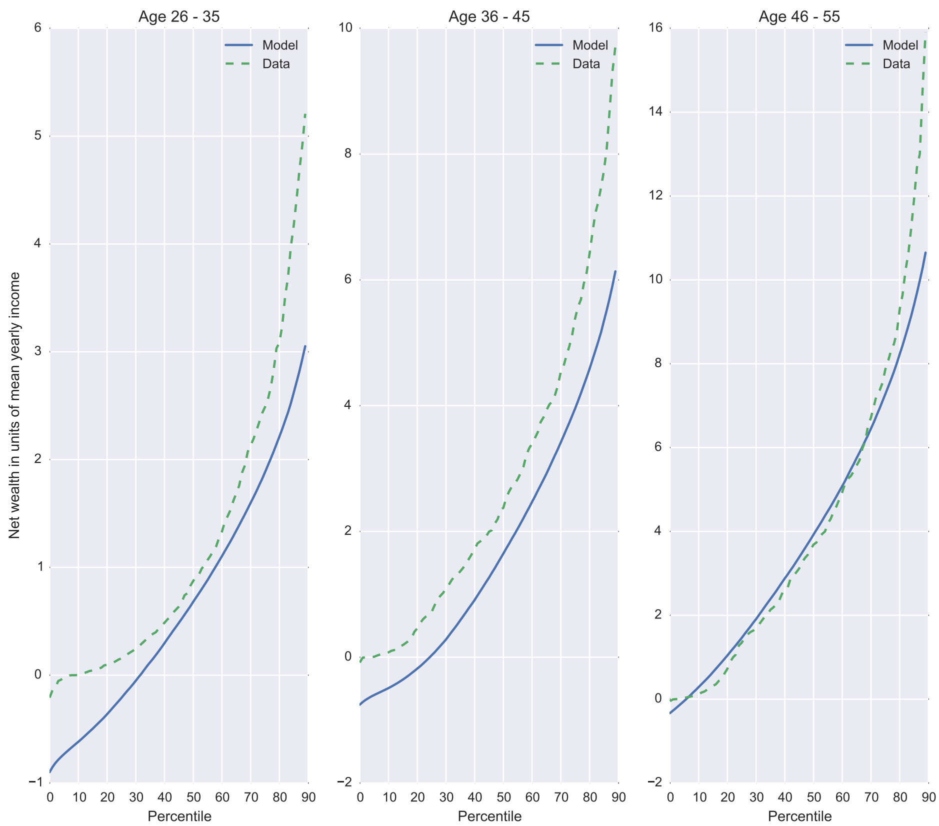


Figure 4.2: Baseline model (PSID HIP income process), calibrated fit by age group



exact income process used in Guvenen (2007). The results show that the estimated discount factor is almost exactly that used in Guvenen's paper, at 0.961, and importantly significantly lower than all other discount factors, estimated including the respective life-cycle profiles of earnings extracted from the different data sets. Intuitively, this result stems from the fact that without the deterministic income profile, the mean of households' expected earnings distributions is much closer to zero, creating a larger precautionary savings motive. The minimum distance estimator then tries to counterbalance this by choosing a lower discount factor, lowering aggregate savings by making households more patient. This suggests that the increase in aggregate savings coming from an HIP based model as reported in the earliest working paper version of Guvenen (2007) is somewhat overstated when the life-cycle profile of earnings is disregarded[21].

4.4 Comparative Statics

Given the largely disappointing results of the calibration and simulation exercises, we now turn to some comparative statics exercises to elicit what features of the model are crucial to get closer to the shape of the observed wealth distribution. To do so, we pick a reasonable baseline calibration from the set of available parameters estimated for different income processes in chapter 3, and then vary each of the 8 parameters governing the model solution by solving the model

[21]Although it has to be noted here that, at least for our measures of gross labour incomes, the processes we estimate might understate the true "disaster risk" in income, given that the sample selection process excluded households with zero earnings. Guvenen et al. (2015) make some effort to alleviating this problem by introducing a mixture of AR(1) components into an HIP process, with one of the AR(1) processes capturing the very low likelihood of extremely large shocks to household income.

Table 4.1: Calibrating the model for different income risk profiles

Income Process	δ	σ
PSID 1968-2013 (no lifecycle)	0.961	1.41
	(0.0008)	(0.006)
PSID 1968-2013 (with lifecycle)	0.973	1.90
	(0.0008)	(0.006)
BHPS gross labour income	0.967	0.621
	(0.00027)	(0.041)
BHPS net labour income	0.965	0.5
	(0.00046)	(0.036)
BHPS net household income	0.975	1.28
	(0.0069)	(0.166)

Table 4.2: Parameters for comparative statics

	δ	σ	ρ	σ_η^2	σ_ε^2	σ_α^2	σ_β^2	$cov(\alpha,\beta)$
Lowest realization	0.94	1.05	0.72	0.01	0.02	0.01	0.00068	-1.
Baseline	0.96	2.0	0.85	0.03	0.05	0.05	0.00038	-0.3
Highest realizatoin	0.98	3.0	0.92	0.05	0.13	0.10	0.00001	0.

in turn for its highest and lowest realization. The parameter values used are summarized in table 4.2.

Changing the variance of the cross-sectional distribution of intercepts does not influence the results in any meaningful ways, as could have been anticipated from the fact that α in effect parallel shifts the entire life-cycle profile of households up or down, which, given that almost all households are far enough away from the borrowing constraint at all times, and in the absence of any different savings behaviour of rich households in the model (as e.g. found in the data by

Dynan et al. 2004), means that savings behaviour is not affected by this change. Similarly, changing the variance of the transitory shock does not alter the results significantly, save for an overall increase in wealth holdings for the highest value of σ_ε^2[22]. The two parameters that have a markedly larger influence on the *shape* of the predicted percentile distribution, and hence help the model get closer to the data moments, are the persistence of the AR(1) component and the variance of its innovations. Figures 4.3 and 4.4 show the effects of varying the persistence of the AR(1) component of the income process for prime age households and households by age group, respectively. When increasing ρ to 0.92, the predicted wealth distribution becomes notably more curved, while the effect of lowering ρ from 0.85 to 0.72 is significantly smaller. This is not very surprising, as the implications of lowering ρ for the half-life of a persistent shock become less severe the lower the starting value of ρ – as figure 4.17 demonstrates, the half life of a persistent shock under the baseline $\rho = 0.85$ is about four years, while for $\rho = 0.72$ it is two years and for $\rho = 0.92$ it is eight years. The model with a high value of the persistent shock performs especially well in capturing the higher wealth accumulation at the higher end of the distribution for the oldest groups of households, which is exactly when we would expect the effect of a series of persistent shocks accumulating over the life-cycle to play the biggest role. Figures 4.5 and 4.6 display the results for changes in the variance of persistence shocks. Just as in the case of an increase in persistence ρ, increasing the variance of the persistent shocks helps to increase the curvature of the predicted wealth distribution, by lowering savings at the lower and and increasing wealth holdings at the upper end at the same time. Indeed, both changes in ρ and in σ_η^2 bring the model parametrisation closer in

[22] Graphical results can be found in the Appendix

line with that of Hintermaier and Koeniger (2011), who are using $\rho = 0.95$ and $\sigma_\eta^2 = 0.47$ in their baseline calibration. Importantly, ρ and σ_η^2 have similar effects on the income distribution that differ from the effects of increases in σ_α^2 and σ_ε^2, as evidenced in table 4.3. It appears that a crucial ingredient if the model if it is to match the empirical wealth distribution is the inequality in lifetime income, and, importantly, the source of this inequality. As can be seen in figures 4.7 and 4.8, changing the dispersion of individual- specific growth rates of income does not have the same effects on the aggregate wealth distribution as changes in ρ or σ_η^2. The reason for this is that rich households in a world in which lifetime income inequality is high mostly because of the size and persistence of permanent shocks need to save in periods of high income, as the effect of the good shock will wear off and might be overlaid by the effects of a large negative shock in the future, while households that are rich in a world where income inequality is high because of inequality in deterministic growth rates will have high income growth across their life-cycle for certain, and hence don't need to save less to achieve consumption smoothing[23]. We then have to conclude that the very essence of the difference between HIP and RIP models of the income process – a lower persistence and variance of the AR(1) component of income, offset by variation in individual-specific, deterministic income growth rates – is what keeps it from matching the empirical profile of wealth holdings. Indeed, our model nests the model in Hintermaier and Koeniger (2011) as a special case with σ_α^2 and σ_β^2 equal to zero, and as figures 4.15 and 4.16 in the appendix show, the model fits the data well with this version of the RIP process. To illustrate the vast improvement in

[23]In fact, to the extent that households know about their high income growth rate early in life, they will save *less* than poor households, who are potentially facing negative income growth rates.

Table 4.3: Standard deviation of lifetime income (multiples of baseline)

	ρ	σ^2_η	σ^2_ε	σ^2_α	σ^2_β	$cov(\alpha, \beta)$
Lowest realisation	-0.09	-0.06	-0.04	0.02	-0.29	-0.22
Highest realisation	0.31	0.35	0.09	0.11	0.41	0.05

model fit, it is instructive to consider the value of the objective function, which gives the sum of the squared difference between model and data moments for all 324 targeted moments: while the best fit of the HIP models implies a value of the objective function of between 350 and 900, the RIP process specified with the parametrisation of Hintermaier and Koeniger (2011) reaches a minimum at 78[24]. The finding that it is mainly the variability of lifetime income that drives wealth accumulation in the model echoes the work of Floden (2008), who shows that the Aiyagari (1994) result of an increase in aggregate wealth holdings in incomplete markets economies with idiosyncratic income variations obtains even when all uncertainty about future income is removed, so that saving is purely driven by the consumption smoothing motive.

4.5 The role of initial beliefs

One question when working with beliefs is obviously how initial beliefs are derived. While the model so far had agents starting with random beliefs centred around their individual-specific true parameter, and we have discussed the role of uncertainty induced by learning, we will now briefly consider situations in which

[24]Our results are not exactly the same as those derived in Hintermaier and Koeniger (2011), as we employ a different solution technique, which implies a less accurate solution to the model, and our model lacks some features present in their paper, notably a rigorous treatment of the US tax system and uncertain lifespans after retirement.

Figure 4.3: Comparative statics for persistence of AR(1) component, prime age

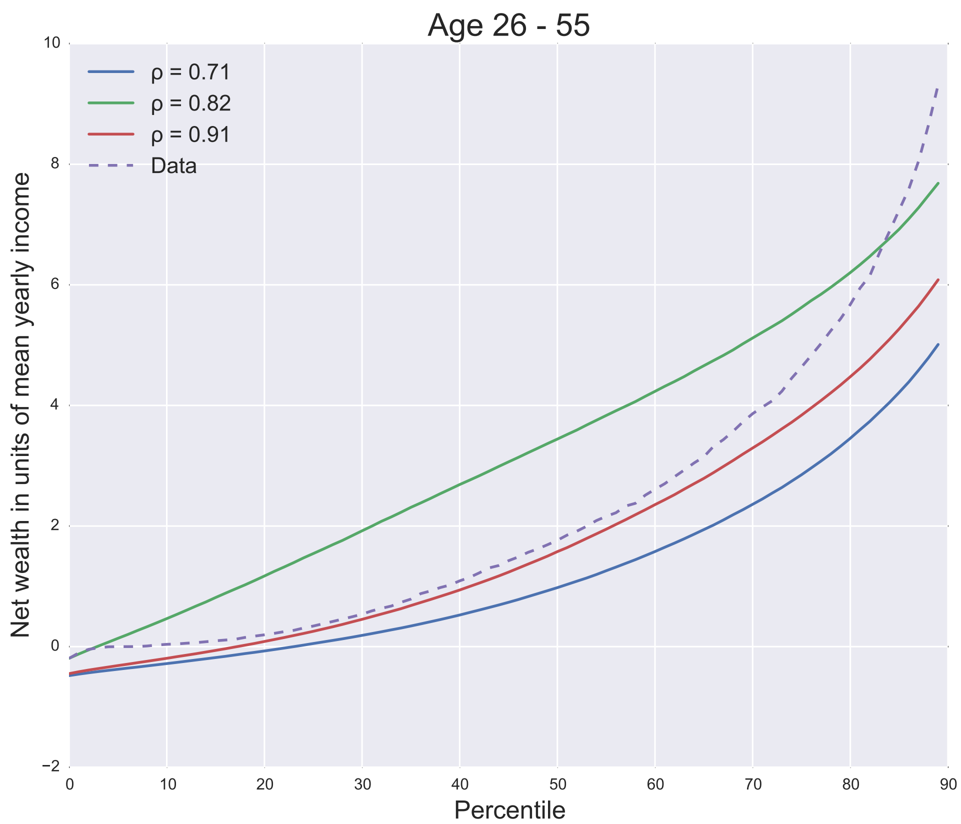


Figure 4.4: Comparative statics for persistence of AR(1) component, by age groups

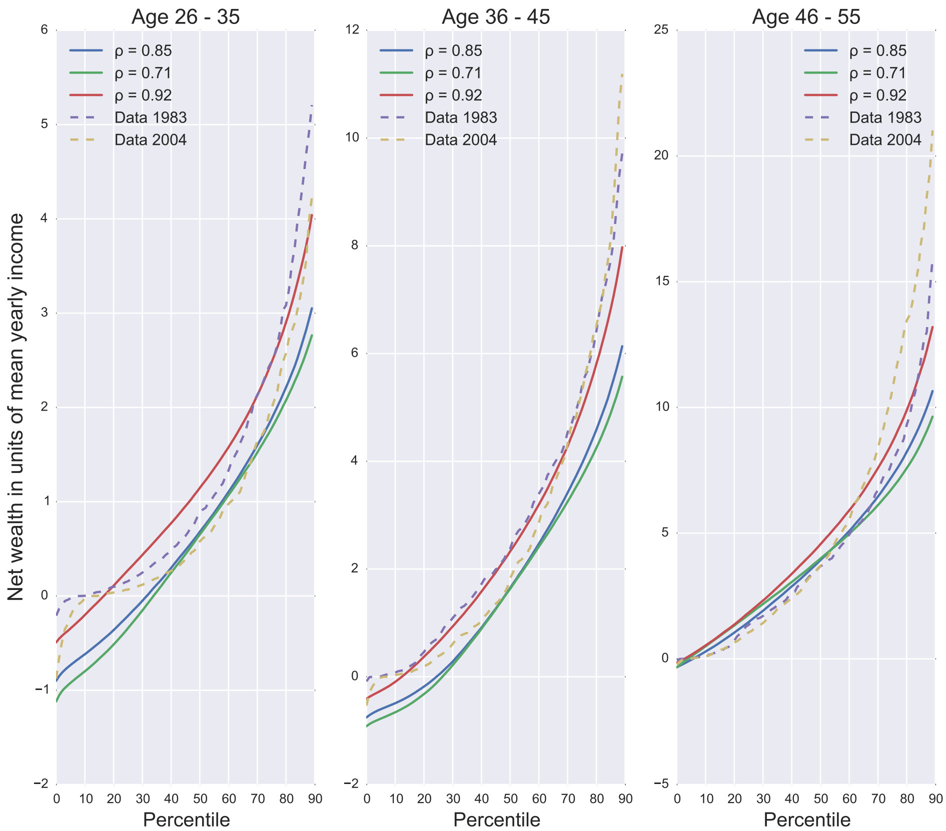


Figure 4.5: Comparative statics for innovation variance of persistent shock, prime age

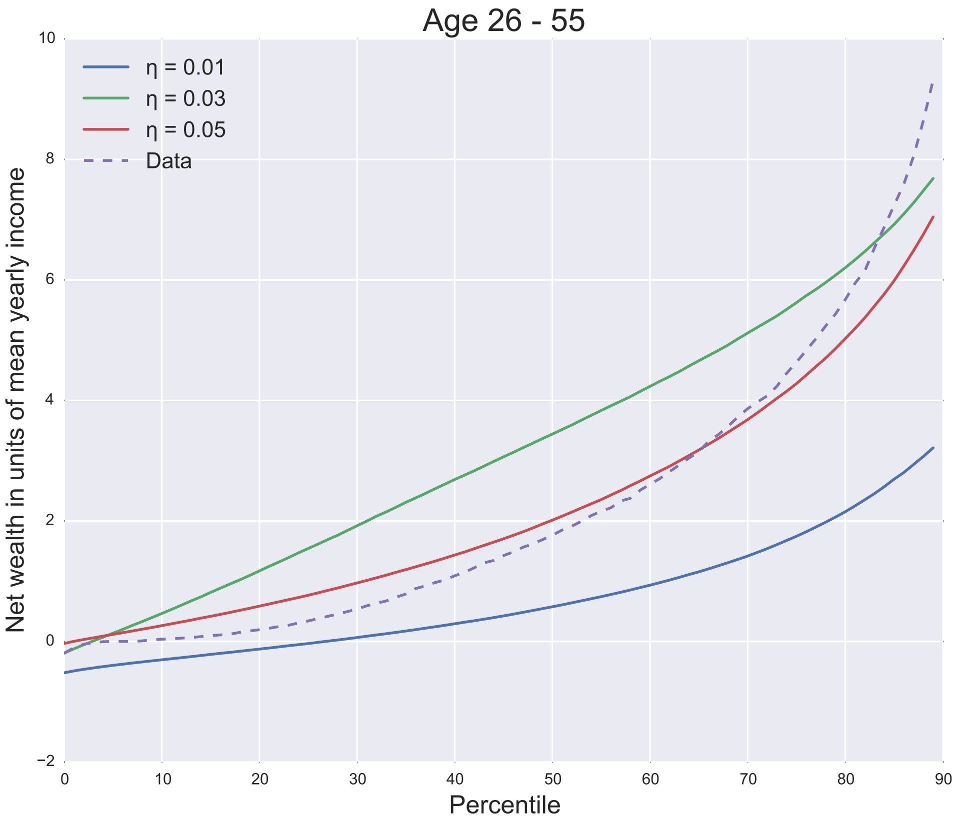


Figure 4.6: Comparative statics for innovation variance of persistent shock, by age groups

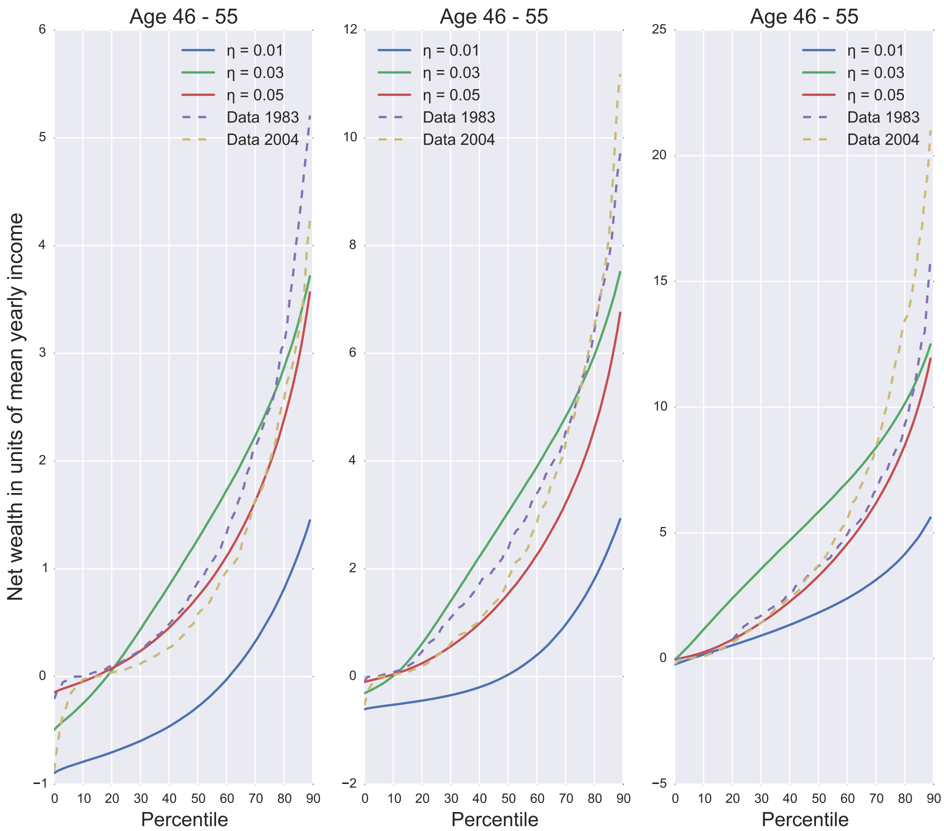


Figure 4.7: Comparative statics for variance of individual-specific growth rates, prime age

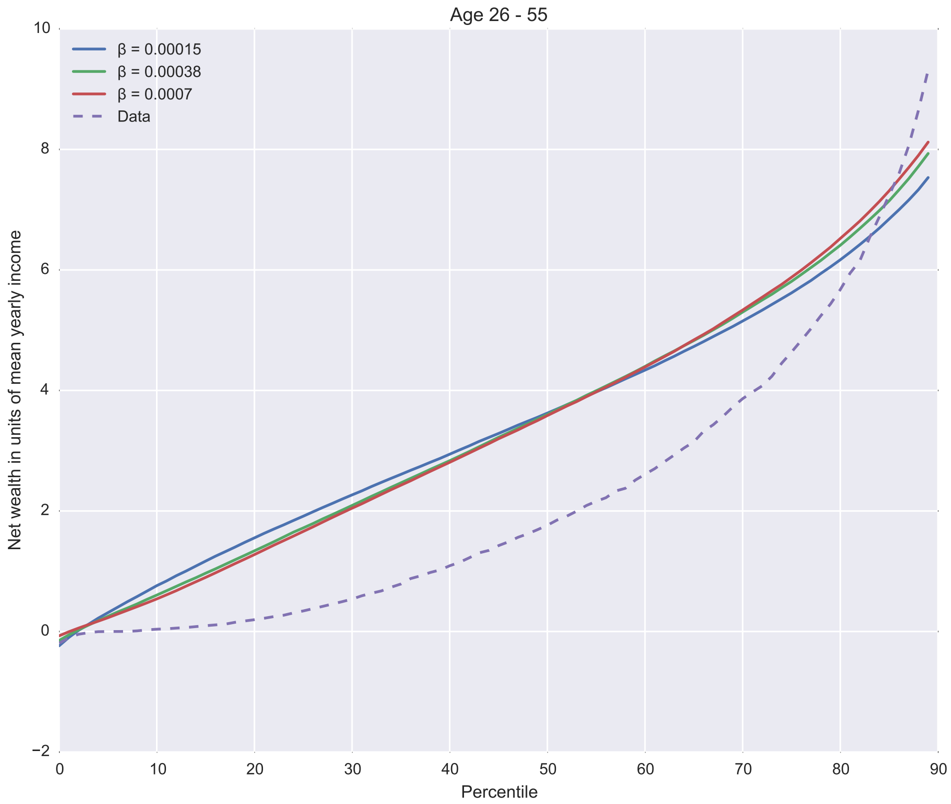
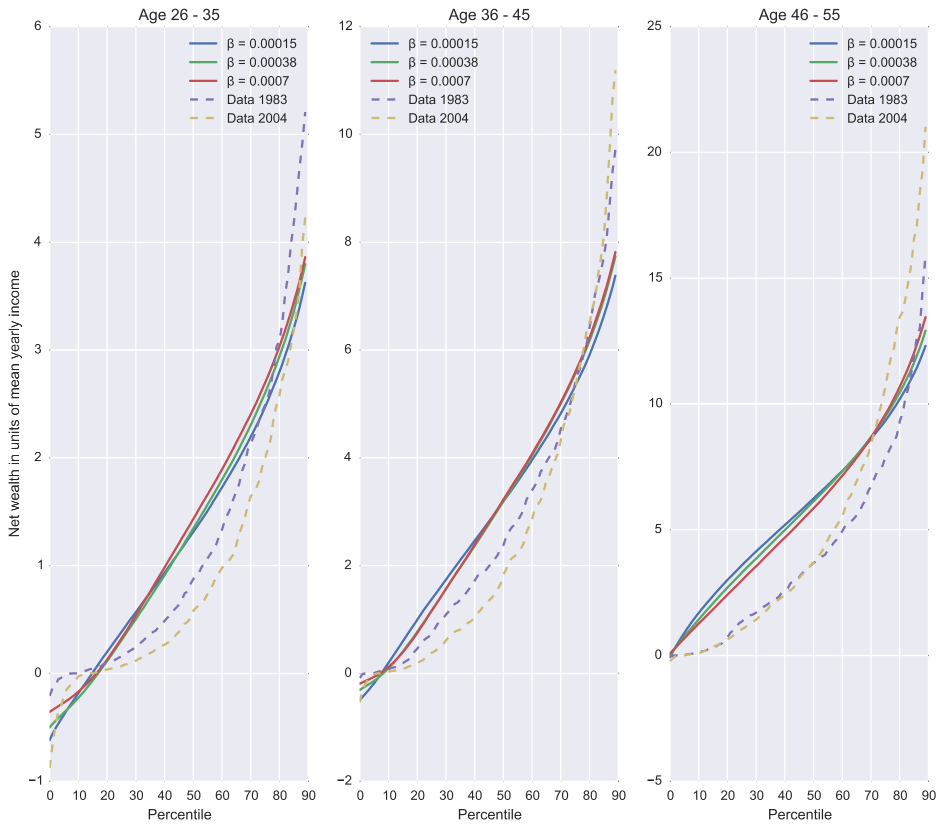


Figure 4.8: Comparative statics for variance of individual-specific growth rates, by age groups



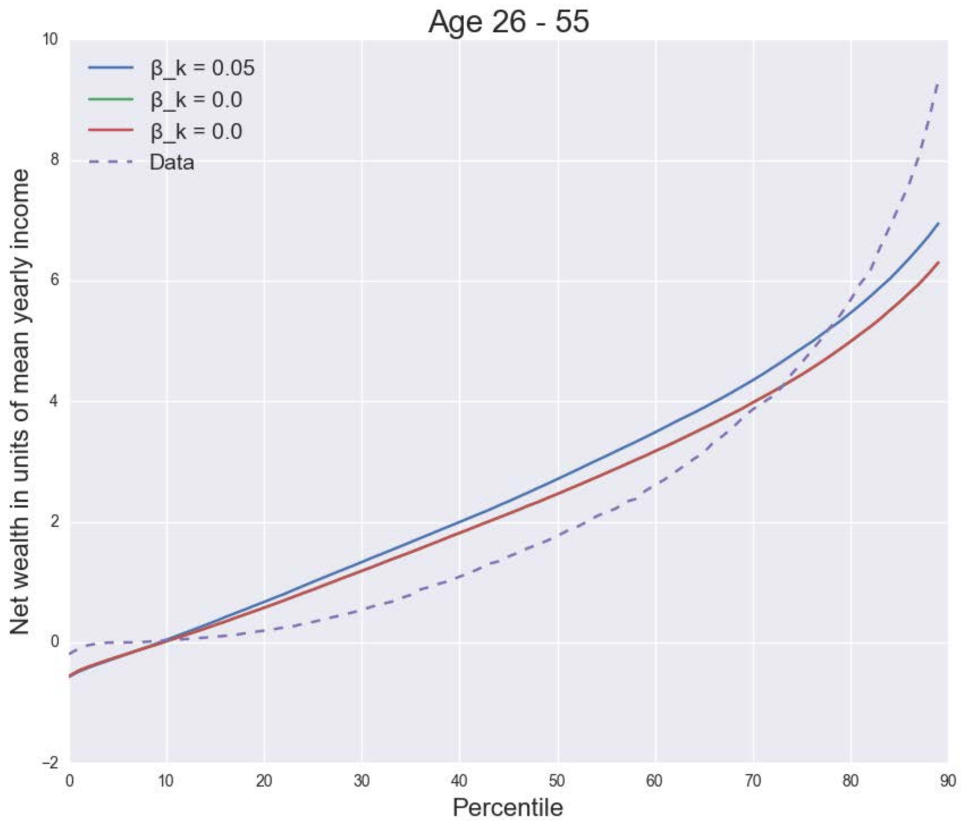
the initial beliefs are systematically incorrect. Of course, with an appropriately chosen set of initial beliefs, the model can be made to produce almost any desired result, so that we will only consider two situations which are at least supported by anecdotal evidence. The first very simple experiment is done under the assumption that agents suffer from overconfidence regarding their own economic fortune. There are a host of studies that offer support for the view that people are too optimistic about their future earnings potential, e.g. a Gallup poll by Moore (2003) in which half of all respondents aged 18 to 29 state that they regard it very or at least somewhat likely to be rich in the future, with the median figure for expected yearly income and wealth cited at $120,000 and $1 million, respectively. Of course it is questionable whether actual economic choices would be based on a vague belief about the indefinite future[25], we can use our model to assess what would happen if people would indeed act on them. Figure 4.9 shows the wealth distribution that obtains if all agents start with a belief that is one percentage point above their original initial belief. The results are not as striking as one might have expected, with the entire wealth distribution apart from the lowest percentile, which is constrained in any case, being shifted down, albeit not by much.

The second scenario under consideration is related to the shifts in the US income distribution happening throughout the 80s and 90s. As has been well documented, wage inequality experienced a secular rise during this period, with top incomes surging, while incomes at the bottom of the distribution saw their growth rates fall. To the extent that these changes were unobserved by agents contemporaneously, and happened through changes in the idiosyncratic growth

[25]Another constraint on this is that these beliefs would require large amounts of borrowing in the present, for which optimistic people would have to find lenders who share their beliefs.

rates, they might have biased beliefs of agents entering the labour market, if they form their beliefs based on past observations of income growth for people in similar places in the wealth distribution. To capture a stylised version of this process in the model we will solve the model assuming that those agents in the upper quintile of the true distribution of β start life with a belief one percentage point below their original initial belief, those in percentiles 60 to 80 half a percentage point lower, those in percentiles 20 to 40 half a percentage point higher, and those in the bottom quintile one percentage point higher. The results in figure 4.10 show a tilting of the wealth distribution, with agents at the upper end accumulating more wealth, since they ascribe a larger part of their good fortune to permanent shocks given that they systematically underestimate their deterministic growth rate. At the same time, more agents at the lower end of the distribution are constrained, as they expect higher future income growth than will actually materialise. These changes in the wealth distribution are consistent with the observed increase in wealth inequality that followed the increase in income inequality in the 1980s and 90s (as discussed in Iacoviello (2008)) and could form the basis of a demand-driven expansion of household debt at the lower end of the income distribution. Indeed, rising income inequality is often cited as a prime reason for the increase in household indebtedness (see e.g. Rajan (2011), Saez and Zucman (2014), or, for a more heterodox treatment, Barba and Pivetti (2009)), although from the standpoint of a standard life-cycle model of consumption and savings this link should be absent, given that recent work based on Social Security records shows virtually the entire increase in income inequality to be due to permanent, rather than transitory shocks, which should result in changes in consumption, rather than wealth inequality. One possibility

Figure 4.9: Effects of optimistic initial beliefs

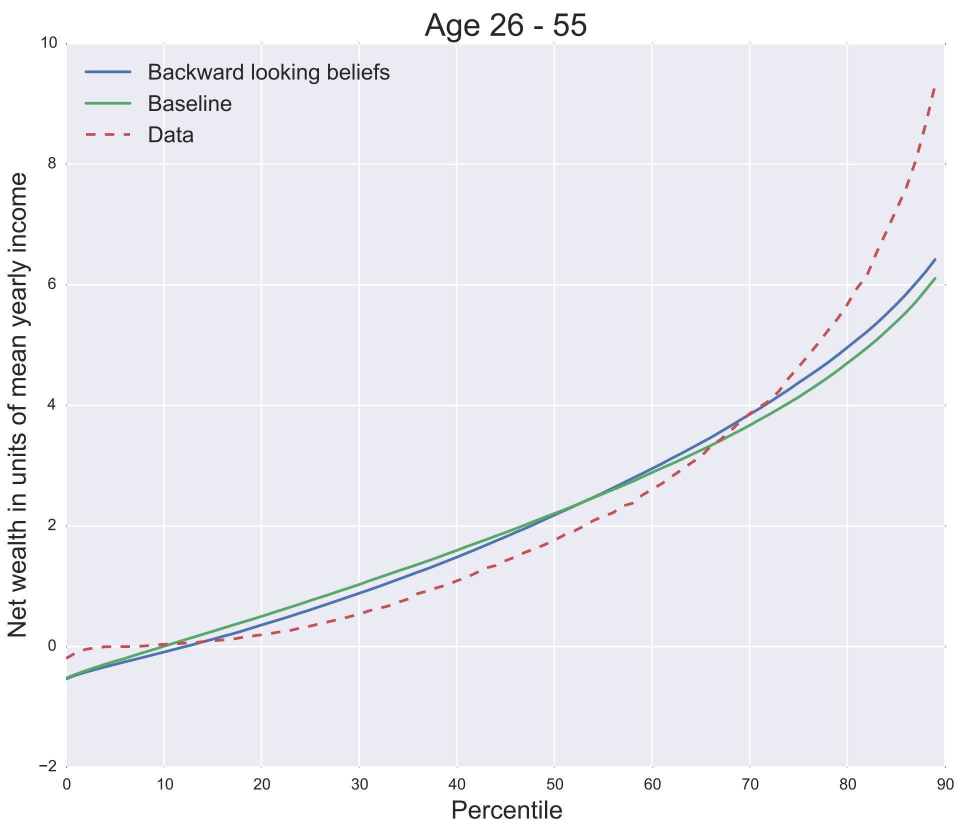


for permanent changes in the income distribution not to immediately filter through to the consumption distribution could then, according to our model, be hysteresis in belief formation of agents entering the labour market at different points of the income distribution.

4.6 Discussion

As this chapter has shown, the learning model of heterogeneous income processes fails in capturing the dynamics of the wealth distribution under all calibrations

Figure 4.10: Effects of beliefs based on previous income growth rates



derived from empirical data on income processes. Comparing the model output of different counterfactual parametrisations, it became clear that the main reason behind this is not the learning mechanism itself, but the different income distribution and risk implied by the heterogeneous income process. To salvage the model, ad-hoc changes to the belief structure of agents can be made, although at this point it becomes a bit of a free-for-all and the model can be made to predict any pattern in the data with a suitable choice of initial beliefs. Building on the work in this chapter, future research should consider the implications of other more realistic income processes on the wealth distribution, to the extent that they can be formulated parsimoniously enough not to increase the computational burden beyond reason. An example would be the work by Meghir and Pistaferri (2004), who model the conditional variance of income shocks using an ARCH model and show analytically that the addition of individual- specific heterogeneity in the innovation variance leads to both a larger dispersion of savings rates and higher aggregate saving. A further case of interest would be the specification derived by Guvenen et al. (2015), which adapts the income process used in this chapter by adding two more AR(1) components with different innovation variances, so that households are subject to potential shocks of different magnitude. As the evidence points to this process being the best description of the income risk households are facing in reality, the implications of this process for the wealth distribution should be investigated further.

4.7 Appendix A: Comparative statics results

Figure 4.11: Comparative statics for variance of individual-specific intercepts, prime age

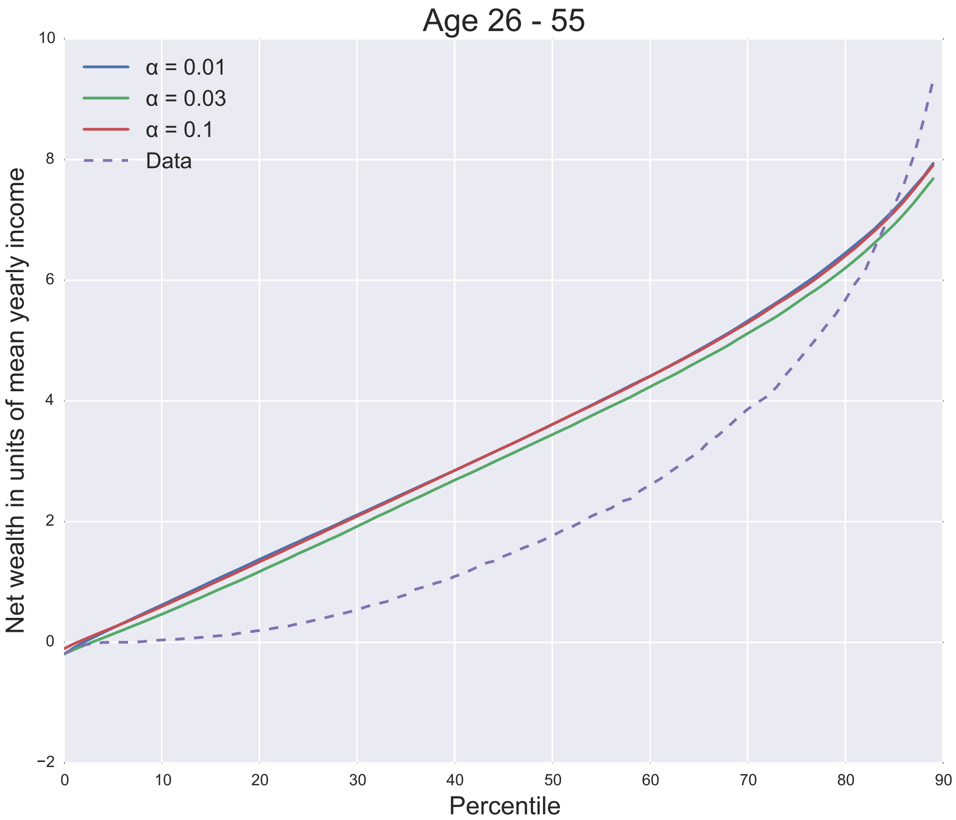


Figure 4.12: Comparative statics for variance of individual-specific intercepts, by age group

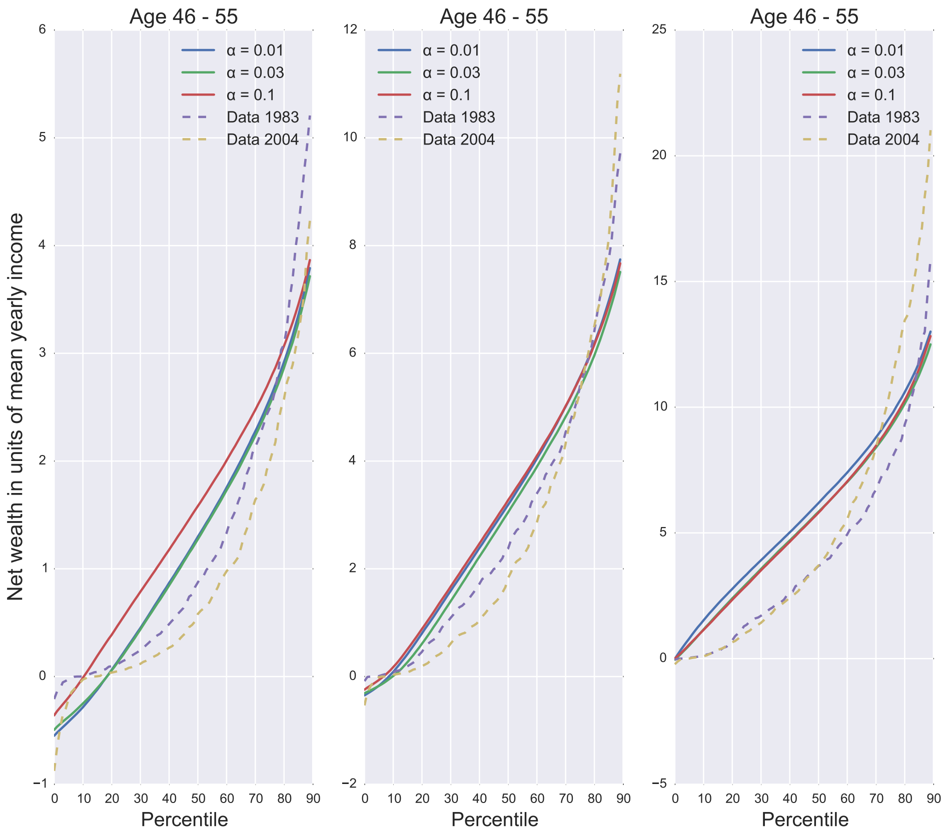


Figure 4.13: Comparative statics for variance of transitory shocks, prime age

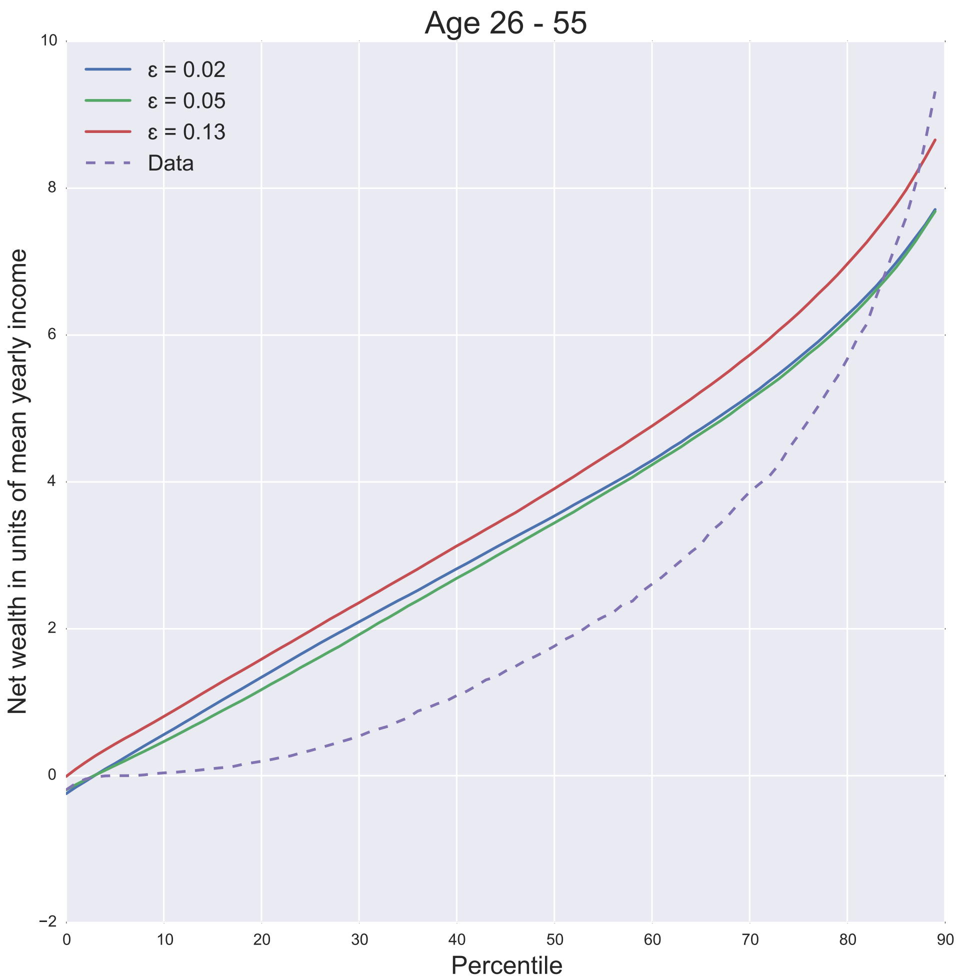


Figure 4.14: Comparative statics for variance of transitory shocks, by age group

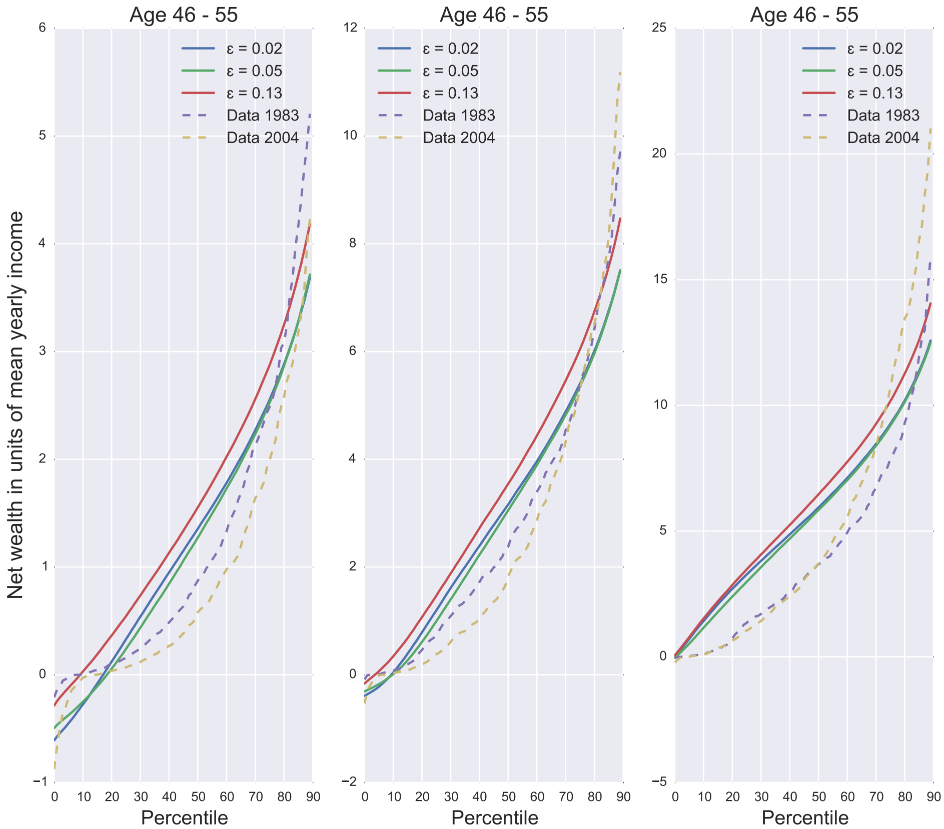


Figure 4.15: Model fit when income is an RIP process with $\rho = 0.95$ and $\sigma_\eta^2 = 0.5$

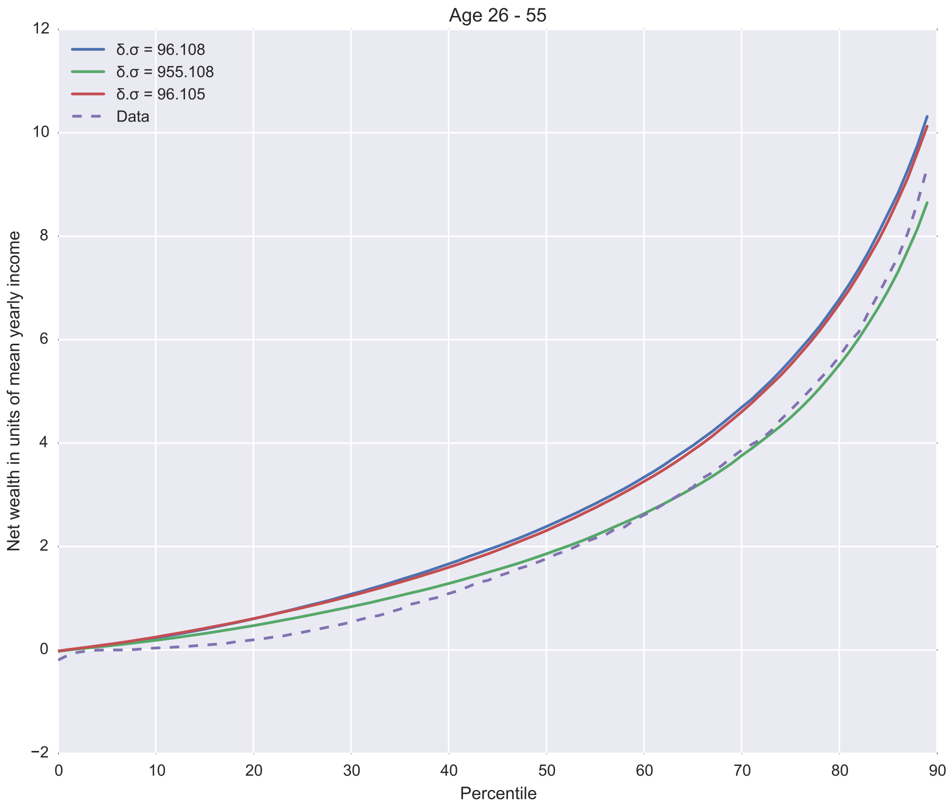


Figure 4.16: Model fit when income is an RIP process with $\rho = 0.95$ and $\sigma_\eta^2 = 0.5$, by age groups

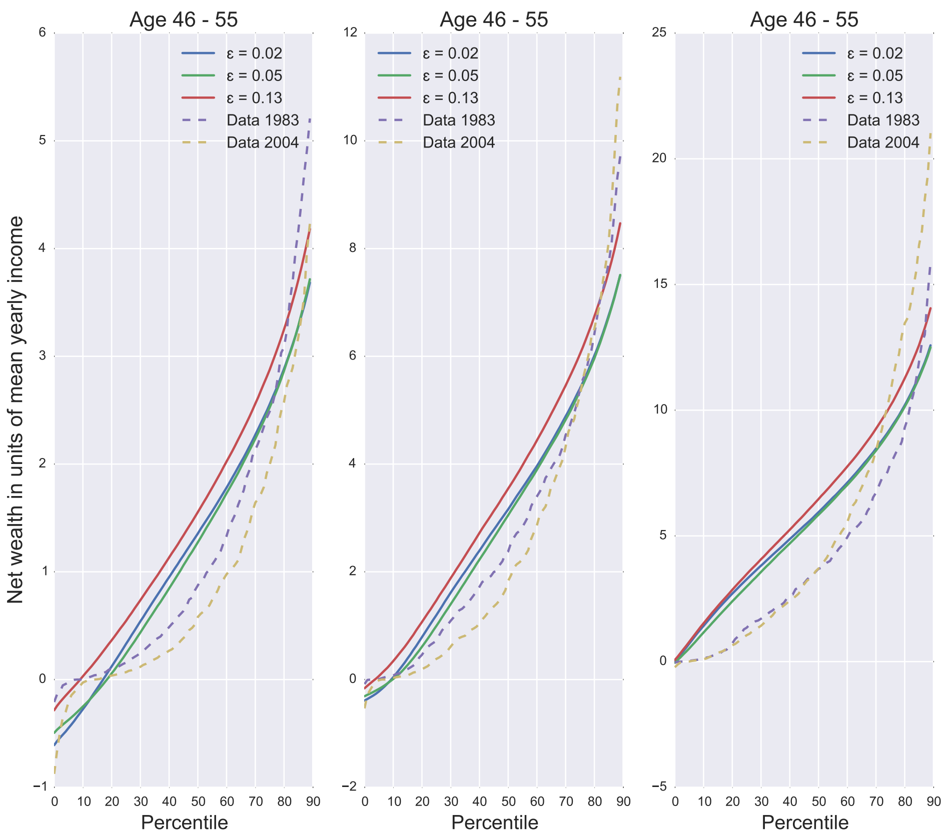
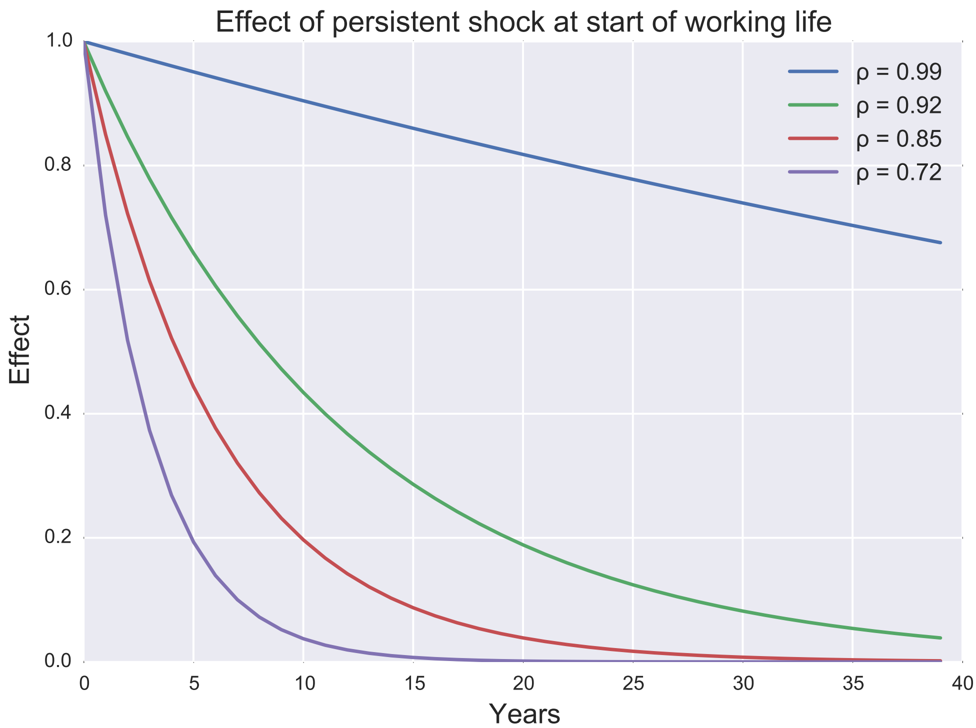


Figure 4.17: Effects of lowering ρ on half life of persistent shocks



Chapter 5

The Competitive Effects of Trade Liberalisation in North America: An Empirical Application of the Melitz and Ottaviano Model

5.1 Introduction

The economic benefits of free trade are arguably one of the most uncontroversial results of economic research, both theoretically and empirically. However, to this date, free trade is by no means uncontroversial in the public sphere, as is evidenced by the fierce opposition that the proposed transatlantic free trade agreement between the US and Europe is facing. Hence, international trade has remained an active field in economic research, a field which has seen major advancements in the past two decades in incorporating firm-level heterogeneity coupled with consumer love of variety into trade models that can account for the firm-level responses to increasing trade openness and the large share of intra-industry trade in the international flow of goods and services [26]. This new vintage of trade models predicts additional welfare gains from trade stemming from reallocations of production to more productive firms (as in Melitz, 2003) or increases in firms' efforts to innovate (as in Grossman and Helpman, 1990). To some extent, these new models of trade also help reconcile the unambiguously positive stance of economic researchers on trade liberalization with the public opposition to it – models taking into account explicitly the heterogeneity across agents of firms within a country show that while on aggregate there are significant efficiency gains from free trade, there are also firms and workers who will lose out individually, and can only benefit from a trade liberalization if either the aggregate gains are redistributed in some way to ensure a Pareto improving allocation, or if they can benefit from the reallocation of production to more productive firms by switching to those firms. DixCarneiro (2014) builds a structural model of the Brazilian labour market to estimate the labour market effects of trade liberalization and finds that, depending on the assumptions about capital mobility, the reallocation of workers across sectors can take up to 30 years.

While these new models of international trade are well grounded in empirical evidence coming from micro data, there are surprisingly few tests of the model predictions for aggregate variables which are decisive for the predicted welfare gains from trade. Recently, Arkolakis et al. (2012a,b) call into question the importance of firm-level heterogeneity by showing that in a class of trade models,

[26]A comprehensive survey of trade models with love of variety preferences and firm-level heterogeneity can be found in Melitz and Trefler (2012).

the additional welfare gains are fairly small and actually even smaller if consumers don't have CES utility. The response of Melitz and Redding (2013) shows that there is still considerable disagreement over how to theoretically evaluate the additional welfare gains from firm selection, and Costinot and Rodriguez-Clare (2014) review the effects of trade liberalizations in a wider class of new trade models to highlight the importance of the market structure under consideration – depending on whether a one- or multi-sector model is used and the degree of competition assumed, gains from trade are estimated to range from 4% to 40% of non-free-trade welfare. These facts motivate us to test the Melitz-Ottaviano model directly in aggregate data on prices, markups and productivity. To do so, we estimate the effects of trade liberalization on the competitive environment in manufacturing markets of the member countries of the North American Free Trade Agreement (NAFTA). We employ an estimation procedure based on the Melitz and Ottaviano (2008) model introduced by Chen et al. (2009), which to our knowledge is the only empirical application of a model with firm-level heterogeneity on aggregate data. Chen et al. (2009) derive estimable regression equations from the model's equilibrium conditions that allow us to test the effects of trade openness on relative price levels, markups and labour productivities of two trading partners. It is further possible to differentiate between the effects of trade in the short run, which, in the model, refers to an economy without relocation decisions for firms, and in the long run, when firms are free to choose their home market for production. However, as the underlying model is static, no direct results on the time path of the impact of trade liberalization can be obtained. We try to address this issue by dividing our sample in ways that make it more amenable to a model-based estimation. Contrary to Chen et al. (2009), we directly observe tariff rates between the three countries in our sample and hence use those as a direct measure of trade openness. Additionally, we test for the effects of third-country trade openness on the relative performance of two countries that are linked through trade, predictions for which can be derived from the multi-country version of the Melitz and Ottaviano model. Our dataset comprises of 64 manufacturing sectors in Canada, Mexico and the US, covering the time period from the introduction of the US-Canadian free trade agreement CUSFTA in 1988 up to 2010, which gives us reason to believe that we are able to capture the long run effects of policy changes even in industries with low firm churning rates.

Our findings support the main model predictions, with tariff barriers stifling domestic competition, leading to higher producer prices and markups as well as lower productivity. In the immediate years after the free-trade agreement when tariff barriers are reduced, relative prices and markups decrease as relative productivity increases, thus giving rise to competitive effects. The results in the long-run, however, are not as clear cut, with some effects reversing as predicted by the model while some effects persist. This is also confirmed by directly looking at the reaction of industries with different entry barriers to changes in trade openness.

The paper is organized as follows: Section 5.2 gives a survey of the previous literature assessing the effects of trade liberalizations in general and of NAFTA specifically. Section 5.3 briefly summarizes the Melitz and Ottaviano (2008) model, derives the most important equilibrium conditions and explains the estimation strategy used in Chen et al. (2009). Section 5.4 then presents our application of the model by giving an overview of the data used and our estimation procedure. The results of our regressions and possible shortcomings as well as extensions of our approach are discussed in Section 5.5; Section 5.6 concludes.

5.2 Related Literature

As free trade has been an active topic in economic research since the times of Ricardo, the literature on the welfare gains from trade is immense. Of particular interest to us of course are papers that investigate the economic effects of NAFTA directly, as well as papers that form the theoretical foundation for our estimation strategy.

The effects of free trade in North America have been scrutinized in a large number of papers over the past two decades, starting with work on the predecessor to NAFTA, the 1987 Canada and US free trade agreement (CUSFTA). Head and Ries (1999) document rationalization effects in Canadian plants as a reaction to decreases in Canadian import duties. Trefler (2004), focusing on the CUSFTA, uses a reduced form econometric approach to find large improvements in labour productivity and decreases in employment after the implementation of CUSFTA, coupled with slightly lower import prices and larger volumes of trade. Fukao, Okubo and Stern (2003) derive regression equations from a partial equilibrium model with imperfect competition to estimate the extent to which NAFTA was trade diverting rather than creating and find responses that vary by industry. Romalis (2007) examines both CUSFTA and NAFTA with a strategy based on estimating demand and supply elasticities and finds a large effect of NAFTA on trade volumes, with only minor price changes and, subsequently, only small changes in welfare. Calderon-Madrid and Voicu (2007) use plant-level panel data from Mexico to show that while productivity increases followed the tariff reductions, the responses of plant-level productivity are very unevenly distributed, with larger plants benefiting disproportionately from productivity increases. The

Melitz (2003) model that is at the heart of our analysis is also put to a test with US manufacturing data by Bernard, Jensen and Schott (2006a), who use plant-level data to estimate the effects of changes in the costs of trade, as measured by tariff rates and transportation cost, on productivity growth and firm entry and exit. Their findings confirm the micro-level implications derived from the assumptions on the productivity distribution in Melitz (2003), which we will highlight in the following section. Other papers have used the structure provided by the Melitz and Ottaviano (2008) model to assess the effects of trade liberalization in other parts of the world: Bellone et al. (2008) use price-cost margins of French manufacturing firms to test the models predictions on the effects of market size, import penetration and exporting status on markups and productivity and confirm that all predictions hold. Corcos et al. (2011) estimate structural parameters in order to simulate counterfactual scenarios by changing the costs of trade between countries. Their exercise shows that the firm selection mechanism is crucial for the magnitude of the welfare gains from trade and the potential gains for a country depend on country size as well as remoteness. The paper that is closest to our own work is Chen et al. (2009), who use the equilibrium expressions for prices, markups and productivity from the Melitz-Ottaviano model to estimate the effects of trade liberalization using a dataset that includes data on 10 manufacturing sectors in seven European countries for the period 1989-1999 with country-pair regressions. There results suggest that trade openness leads to an increase in competitiveness in the short-run with diminishing and at times reversed effects in the long-run, as predicted by the model.

5.3 Model and Estimation Equations

The Melitz and Ottaviano (2008) model is a synthesis of the contributions of Melitz (2003), who introduces firm heterogeneity through random draws of a cost parameter for firms entering the market, and Ottaviano et al. (2002), who develop a model with endogenous markups arising from a linear consumer demand system with horizontal product differentiation. The model yields equilibrium conditions that determine a cost cut-off level, i.e. a level of productivity below which firms are not able to compete in the marketplace. This cut-off level uniquely determines all relevant aggregate variables in the model, namely the distribution of prices, markups and productivity. Importantly, the equilibrium conditions of the model economy are different depending on whether firm entry is allowed or not. Without firm entry, the model captures a short-run equilibrium, with the cost cutoffs in two markets given by:

$$N = \bar{N}\left(\frac{c_D}{c_M}\right)^k + \bar{N}^* \frac{1}{\tau^k}\left(\frac{c_D}{c_M^*}\right)^k \tag{5.1}$$

$$N^* = \bar{N}^*\left(\frac{c_D^*}{c_M^*}\right)^k + \bar{N} \frac{1}{(\tau^*)^k}\left(\frac{c_D^*}{c_M}\right)^k \tag{5.2}$$

Here, a star denotes the foreign market, $\bar{N}$ is the fixed number of incumbents in a market and N is the number of firms that are producing. c_M is the upper bound of the distribution of cost draws, c_D is the cut-off level, i.e. the highest cost draw that allows a firm to earn non-negative profits. $\tau > 1$ is the iceberg cost of trade faced by foreign companies exporting to the domestic market and can be interpreted as a measure of trade costs, tariffs and other impediments to trade.

The long-run equilibrium of the economy allows for firm entry into a market, so that the number of firms in a market is now endogenously determined by a zero profit condition for entrants that balances a fixed cost of entry with the expected profits when drawing a cost level from the (known) cost distribution of a country. The equilibrium conditions pinning down the cost cut-off are

$$c_D = \left[\frac{\phi c_M^k}{L} \frac{1-(\tau^*)^{-k}}{1-(\tau\tau^*)^{-k}}\right]^{\frac{1}{k+2}} \tag{5.3}$$

$$c_D^* = \left[\frac{\phi c_M^k}{L^*} \frac{1-\tau^{-k}}{1-(\tau\tau^*)^{-k}}\right]^{\frac{1}{k+2}}, \tag{5.4}$$

where L is the size of the domestic market. Since all aggregate variables in the Melitz and Ottaviano model are linear functions of the cost cut-off, equations describing the relative price, markup and productivity levels in two countries connected by trade can easily be found by simply dividing the expressions for c_D by those for c_D^*. This gives, for the price level in the short run:

$$\left(\frac{\bar{p}}{\bar{p}^*}\right)^k = \left(\frac{c_D}{c_D^*}\right)^k = \left(\frac{c_M}{c_M^*}\right)^k \frac{\bar{N}^*}{\bar{N}} \frac{N}{N^*} \frac{1 + \frac{\bar{N}}{\bar{N}^*} \frac{1}{(\tau^*)^k} \left(\frac{c_M^*}{c_M}\right)^k}{1 + \frac{\bar{N}^*}{\bar{N}} \frac{1}{\tau^k} \left(\frac{c_M}{c_M^*}\right)^k} \tag{5.5}$$

and in the long run:

$$\left(\frac{\bar{p}}{\bar{p}^*}\right)^{(k+2)} = \left(\frac{c_D}{c_D^*}\right)^{(k+2)} = \left(\frac{c_M}{c_M^*}\right)^k \frac{L^*}{L} \frac{1-\frac{1}{\tau^k}}{1-\frac{1}{(\tau^*)^k}} \tag{5.6}$$

These two equations capture one of the central predictions of the Melitz and Ottaviano model: asymmetrical trade liberalizations will have opposing effects on competitiveness in the short and the long run. By equation (5.5), lowering trade barriers induces a fall in the cost cutoff, and hence decreases in prices and markups and increases in productivity. In the long run, however, the effects are reversed, as an increase in trade costs induces firms to choose the relatively more protected market for production, thereby increasing competition in markets that are shielded from foreign firms.

Chen et al. (2009) show that it is possible to substitute out the trade cost term with an openness term that is derived from a measure of foreign firms market share in the domestic market. However, since we are interested in the effect of tariff rates on competitiveness, we use tariff data directly as a proxy for τ. This strategy should pick up the effects of tariff rates in our estimation if other determinants of trade openness (e.g. oil prices (Kilian et al., 2009), credit conditions (Chor and Manova, 2012), shared culture and language between countries) do not vary systematically across industries. However, as a first step, we will replicate their analysis exactly in our data set (albeit with different instruments for openness), which requires us to make the same substitution, which is:

$$\frac{1}{\tau^k}\left(\frac{c_M}{c_M^*}\right)^k = \frac{\theta}{1-\theta} \tag{5.7}$$

Similarly, an expression for the average markup can be derived. The determination of the average markup is equivalent to the one for average prices so expressions for the short– and long–run impacts of openness on markups can readily be derived. Somewhat more problematic is the index for productivity, as the model requires knowledge of a firm's unit costs c, which are not observable. Chen et al. work around this issue by assuming away differences in capital costs, so that average industry productivity can be approximated by the ratio of nominal wages to labour productivity: $\bar{c} = \frac{w}{z}$. If it is additionally assumed that unit labour costs only depend on nominal wages, the ratio of domestic to foreign labour productivity can be written as:

$$\frac{z}{z^*} = \frac{w}{w^*}\frac{\overline{c}^*}{\overline{c}} \tag{5.8}$$

If the least competitive firm in an industry with a productivity draw at the upper bound of the distribution c_M has labour productivity z_M and labour is perfectly mobile between firms, equation (5.8) implies $\frac{z}{z^*} = \frac{w}{w^*}\frac{c_M^*}{c_M}$. This relationship can then be used in an analogous fashion as before to construct an expression relating openness to productivity. In the short run, equation (5.8) can be amended to yield:

$$\left(\frac{z}{z^*}\right)^k = \left(\frac{z_M}{z_M^*}\right)^k \frac{(\bar{N}/N)}{(\bar{N}^*/N^*)} \frac{1+\frac{\bar{N}*}{\bar{N}}\frac{\theta}{1-\theta}}{1+\frac{\bar{N}}{\bar{N}^*}\frac{\theta^*}{1-\theta^*}} \tag{5.9}$$

Higher values of θ thus lead to higher productivity (conditional on $\bar{N}/N$), as they force lower productivity firms to shut down production. For the long run, equation (5.5) combined with the expression for labour productivity gives:

$$\left(\frac{z}{z*}\right)^{k+2} = \left(\frac{w}{w^*}\right)^2 \frac{L}{L^*}\left(\frac{z_M}{z_M^*}\right)^k \frac{1-\frac{\theta}{1-\theta}}{1-\frac{\theta^*}{1-\theta^*}} \tag{5.10}$$

Larger markets exhibit higher labour productivity, while the effects of θ and θ^* are the opposite of those in the short–run.

5.3.1 The Role of Market Entry

Following the arguments in Trefler (2004), we want to exploit the nature of NAFTA being close to a natural experiment and hence try to identify the effect of the policy measures (i.e. the changes in tariff rates) separately from the effects of trade openness in general. There is evidence that trade openness – measured by the import penetration of a certain country or industry, as in Chen et al. (2009) – is affected by a number of external forces, including oil prices (Kilian et al., 2009), credit conditions (Chor and Manova, 2012), shared culture and language between countries and many more. Therefore, we deconstruct the iceberg costs of trade into two parts: $\tau^{lh} = \frac{Tr^{lh}}{\theta^{lh}}$, where $Tr^{lh} > 1$ is the tariff rate for trade between countries l and h and $\theta^{lh} > 1$ captures the additional costs of trade imposed by the aforementioned factors. Then, the import penetration can be viewed as a measure of θ^{lh}, as in Chen et al. (2009), while a carefully constructed tariff measure can directly account for the effect of NAFTA-mandated changes in the trade environment. Obviously, the import penetration in a sector depends on the tariff measure as well, so an instrumental variable approach has to be taken in order to identify this effect. We defer the discussion of the construction of the tariff measure and the choice of instruments to the following chapter. Prior work of Bernard et al. (2006b) suggests that tariff rates throughout the 1980s, at an average level between four and five percent, accounted for about the same fraction of trade costs as costs directly attached to shipping the good, i.e. freight and insurance, so we expect them to have a sizeable impact on trade flows between countries and hence the competitive environment.

As we have seen in the exposition of the Melitz and Ottaviano model above, there is one crucial caveat in taking the model to the data: due to the static nature of the model, the comparative static results just compare one steady state with another, while being silent about the transitional dynamics. The estimation strategy of Chen et al. (2009) tries to account for this by estimating an error correction model to identify the long-run separately from the short-run, but their results – just as ours – are mixed for the long run and it cannot be ruled out that this is due to the estimation procedure. Therefore, we try to address this issue in a more direct way: as short- and long-run in the model differ only in the possibility of firm entry, we separate industries into those with a fixed number of firms and those with low entry barriers. This distinction then gives us industries that represent the short- and long-run and we can directly investigate whether the coefficients on the relevant variables differ significantly[27]. This approach, however, leads to two issues that need to be addressed before implementation. First, it is not *a priori* obvious how to measure the entry conditions in an industry; while the theoretical model uses the number of firms, this could in practice either refer to firms or to establishments (i.e. different production sites run by the same parent company), or even to employees, as firms in the model use unit labour input. Second, there is no reason to believe that different measures of entry and exit dynamics are exogenous with respect to trade openness – indeed in the model trade openness is a key factor in the entry decision of firms, but in the real world there might be various other factors that might lead to industries being asymmetrically affected by a change in trade costs, hence biasing our results. To tackle both these issues, we aim to construct a robust measure of industry dynamics by aggregating multiple studies that examine firm and employment turnover in Canada, Mexico and the United States as well as Europe over different

[27]We were inspired to do so by Head and Ries (1999) who use the classification to test competing theories of trade that rely on different market structures.

time periods. With this, we hope to identify those industries that are either very dynamic or very static over a broad set of different measures, regions and time periods. Table 5.1 gives an overview of the studies used and a glance at their respective results, showing considerable variation in the dynamics of entry and job creation in different manufacturing sectors.

In order to aggregate the different studies, we compute percentile-based rankings of the industries for each study (to account for the different number of industries across studies) and then average the percentiles across studies. Based on these average percentiles, we can then split the sample according to the short- and long-run distinction made in the model: those industries above the 70th percentile are taken to represent the dynamic, ”free entry” sample and thus the long run, while those industries below the 40th percentile are taken to represent the short run. This procedure leads us to split the sample three-ways: Tobacco, Food Processing, Paper, Chemicals, Primary Metals and Petroleum industries fall into the long run category, while Furniture, Wood, Non-electrical Machineries, Fabricated Metals, Printing, Apparel, and Instruments are taken to represent the short run of the model. The remaining industries are too close to the median to be classified either way and are thus dropped from the sample, which leaves us with 1863 year-industry-country pair observations for the free entry sample, and 1701 observations for the fixed entry sample[28].

A little thought experiment may clarify the role that market entry effects play

[28]Due to different classification systems, the aggregation of studies was not always exact and some industry groups are quite heterogeneous when sub-industries are considered. For further details on the aggregation see Appendix B

Table 5.1: Market Structure measures used, numbers in percent

Study	Subject	Highest Turnover	Lowest Turnover
Dunne et al. (1988)	Entry Rates (4-yearly) U.S. 63-82	Instruments (60.3) Lumber (49.70) Printing (49.0)	Leather (29.4) Food Processing (23.9) Tobacco (20.5)
Samianego (2008)	Entry Rates (yearly) Europe 97-04	Paper, printing, software (15.6) Textiles (11.9) Petroleum and Coal (11.9)	Chemicals (9.5) Plastics (9.4) Food Products (9.1)
Brown (2004)	Employment renewal Canada 73-96	Plastic (79.5) Furniture (79.4) Fabricated Metals (77.2)	Primary Metals (33.6) Paper (32.4) Tobacco (4.2)
Foster et al. (2006)	Job creation (yearly) U.S. 72-98	Lumber (11.8) Apparel (11.2) Miscellaneous (11.0)	Paper (5.9) Petroleum (5.9) Tobacco (5.1)
Baldwin et al. (1994)	Job turnover (yearly) Canada 73-86	Furniture (26.5) Machinery (26.3) Lumber (25.7)	Petroleum (14.1) Primary Metals (13.5) Paper (10.7)
Baldwin et al. (1994)	Job turnover (yearly) U.S. 73-86	Lumber (27.2) Apparel (25.5) Leather (22.5)	Petroleum (14.6) Chemicals (14.0) Paper (13.3)

in muddling the distinction between short– and long–run equilibria. The Melitz and Ottaviano (2008) model yields opposing predictions on the effects of trade liberalization on country-level economic variables such as prices, productivity and mark-ups. The reason for the differences, as we have seen, lies in the assumptions on market structure: there are two different equilibria depending on whether entry into a market is allowed. We repeat them here for convenience:

$$c_D^k = c_M^k \frac{\bar{N}^*}{N^*}\left(1 + \frac{\bar{N}}{\bar{N}^*}\frac{\theta^*}{1-\theta^*}\right)$$
$$c_D^{k+2} = \frac{\phi c_M^k}{\Upsilon L}\left(1 - \frac{\theta^*}{1-\theta^*}\right)$$

In model terms, only one of these two equations holds at any given time, and it is posited that the first equation captures the short run effects of trade liberalization, while in the long run firms are allowed to enter the markets and the effects of trade barriers are determined by the second equation. No further assumptions on the nature of the firm's entry decisions or capital adjustment costs are made that could help separate short- from long run. However, in reality, it seems to be more natural to assume that there is a gradual evolution from one equilibrium to the other, and this view is borne out by data on firm entry and exits showing that in a given year, only between five and ten percent of firms in a given industry are new entrants, while over longer horizons this figure goes up to 80 percent. Hence, it seems to be reasonable to model the transition from the short- to the long run equilibrium by introducing a parameter α that governs the fraction of firms entering an industry. The effects of this parameter are most clear on the productivity side, given that firms cannot change their productivity level, the new productivity distribution will be a weighted average of new entrants' and existing firms' productivity. As the examples in Chen et al. are formulated with respect to relative prices, and we are using their notation, we will discuss the effects of limited firm entry in the price level case as well. The argument carries through if one is ready to assume a nominal rigidity that prevents incumbents from re-optimizing their prices, similar to the assumptions made in New Keynesian monetary models. Similar to the productivity level, the price level is then a weighted average of new and old prices (for simplicity, here we abstract from substitution effects induced by the new relative prices of new and old producers):

$$
\begin{aligned}
\bar{p} &= \alpha \bar{p}^{LR} + (1-\alpha)\bar{p}^{SR} \\
&= \alpha c_D^{LR} + (1-\alpha) c_D^{SR} \\
&= \alpha \left(\frac{\phi c_M^k}{\Upsilon L} \left(1 - \frac{\theta^*}{1-\theta^*} \right) \right)^{\frac{1}{k+2}} + (1-\alpha) \left(c_M^k \frac{1}{\frac{\bar{N}}{N} \left(1 + \frac{\bar{N}^*}{\bar{N}} \frac{\theta}{1-\theta} \right)} \right)^{\frac{1}{k}}
\end{aligned}
$$

where the second line drops the constant linking price level and cost cut-off for notational simplicity. It can easily be seen that the introduction of the α parameter makes the expression for the price level hugely complicated and eliminates the possibility to cancel out most constant terms by using relative prices as was done in Chen et al. (2009). Obviously, the above expression is impossible to take to the data in the hope of identifying any of the parameters.

Let's consider a simplified version of the above. Assume that relative prices levels in the short- and long run, respectively, are given by:

$$\frac{\bar{p}^{SR}}{\bar{p}^{*SR}} = \left(\frac{c_M}{c_M^*}\right)^k \frac{(\bar{N}^*/N^*)}{(\bar{N}/N)} \frac{\rho^*}{\rho}$$
$$\frac{\bar{p}^{LR}}{\bar{p}^{*LR}} = \left(\frac{c_M}{c_M^*}\right)^k \frac{L^*}{L} \frac{(1-\rho^*)}{(1-\rho)}$$

This is a simplified version of the equilibrium conditions in Chen et al. using the notation of Melitz and Ottaviano in which trade freeness is measured by $\rho \in (0, 1)$. It captures the main essence of the model, in the short run relative prices depend on the number of firms and negatively on trade freeness (increasing ρ will decrease $\bar{p}$), while in the long country size matters and prices depend positively on trade freeness (increasing ρ decreases $1 - \rho$ and thus increases $\bar{p}$). Now assume further, that price setting decisions and substitution behaviour of consumer is such that we can aggregate *relative* price levels in the same way we aggregated individual price levels before. Then:

$$\begin{aligned}\frac{\bar{p}}{\bar{p}^*} &= \alpha \frac{\bar{p}^{LR}}{\bar{p}^{*LR}} + (1-\alpha)\frac{\bar{p}^{SR}}{\bar{p}^{*SR}} \\ &= \alpha \left(\left(\frac{c_M}{c_M^*}\right)^k \frac{L^*}{L} \frac{(1-\rho^*)}{(1-\rho)} \right) + (1-\alpha) \left(\left(\frac{c_M}{c_M^*}\right)^k \frac{(\bar{N}^*/N^*)}{(\bar{N}/N)} \frac{\rho^*}{\rho} \right)\end{aligned}$$

Here, the fundamental identification problem becomes apparent: in the first term on the right hand side of the equation, the effect of ρ on $\bar{p}$ is positive, while in the second term it is negative. However, the size and sign of the composite effect will be governed by α, which is unobservable. In order to estimate the effects of trade openness on prices, we have to control for firms entry behaviour. While this might well be endogenous to changes in trade policy, it is reasonable to assume that different industries have different entry conditions due to fixed costs inherent in the business model. We can try to exploit this variation in entry conditions by sorting businesses according to the ease of entry; then, ceteris paribus, an industry with lower barriers to entry should exhibit a response to trade liberalization along the lines that the model predicts for the long run equilibrium (as the value of α increases, $\bar{p}$ approaches $\bar{p}^{LR}$), while an industry with high entry barriers subject to the same trade liberalization should see a very different reaction.

One way to alleviate this problem is by trying to use information on α in the estimation. Splitting the sample based on our aggregated turnover measures can be seen as a crude approximation to this, as can be the construction of dummy variables for high and low turnover industries. The most direct way, however, would be to use information on industry turnover rates directly. Obviously, this brings back the very same endogeneity problems we described above that were one reason to aggregate the studies in the first place, which makes it important to instrument for entry and exit rates in industries using turnover measurements for different periods than the one considered in the estimation. The variable construction will be explained in more detail in the next section.

5.4 Application

Starting from the equations for prices, productivity and markups derived within the Melitz-Ottaviano framework, we can derive estimable log-linearised equations analogous to those in Chen et al. The estimation equation for prices is given by:

$$\begin{aligned}\Delta \ln \left(\frac{\bar{p}_{it}}{\bar{p}^*_{it}}\right) =& \beta_0 + \beta_1 \Delta \ln \tau_{it} + \beta_2 \Delta \ln \tau^*_{it} + \beta_3 \Delta \ln D_{it} + \beta_4 \Delta \ln D^*_{it} \\ &+ \gamma \left[\ln \left(\frac{\bar{p}_{it-1}}{\bar{p}^*_{it-1}}\right) + \delta_0 + \delta_1 \ln L_{t-1} + \delta_2 \ln L^*_{t-1} + \delta_3 \ln \tau_{i,t-1} + \delta_4 \ln \tau^*_{i,t-1}\right] + \varepsilon_{ijt}\end{aligned} \tag{5.11}$$

In the above equation, the number of firms serving the domestic market, N, has been replaced by the more readily observable number of domestic firms producing for the domestic market, D, where $D = N\left(\frac{c_D}{c_M}\right)^k$. The short–run dynamics are estimated in the first part of the equation, with regressors expressed in first differences. The long run is represented by the term in brackets. From the perspective of this model, we would expect $\beta_1 > 0$, an increase in domestic import tariffs increases relative prices in the short-run, and correspondingly $\beta_2 > 0$. The model predicts a dampening effect of the number of domestic firms on domestic prices, which should be reflected by $\beta_3 < 0$, and the opposite for foreign firms, $\beta_4 < 0$. As we expect the coefficient on the error correction term, γ, to be negative, a reversal the

As previously discussed, all aggregate variables (prices, markups, productivity) are ultimately functions of the cost-cutoff level c_D, leading to a very similar estimation equations for our other dependent variable. The effect of tariffs, openness, number of firms and market size on productivity is estimated by:

$$
\begin{aligned}
\Delta \ln\left(\frac{\bar{z}_{it}}{\bar{z}^*_{it}}\right) =& \beta_0 + \beta_1 \Delta\tau_{it} + \beta_2 \Delta\tau^*_{it} + \beta_3 \Delta \ln D_{it} + \beta_4 \Delta \ln D^*_{it} \\
& + \gamma \left[\ln\left(\frac{z_{it-1}}{z^*_{it-1}}\right) + \delta_0 + \delta_1 \ln L_{t-1} + \delta_2 \ln L^*_{t-1} + \delta_3 \tau_{i,t-1} + \delta_4 \tau^*_{i,t-1} \right. \\
& \left. + \delta_5 \ln w_{i,t-1} + \delta_6 \ln w^*_{i,t-1} \right] + \varepsilon_{ijt}
\end{aligned}
\tag{5.12}
$$

where δ_7 and δ_8 capture the effects of changes in nominal wages in the long run. The intercepts β_0 are introduced to capture differences in country–specific technology as Chen et al. depart from the baseline Melitz-Ottaviano model by allowing for such differences. While in the baseline, these vary by country-pair, we check the robustness of the specification by allowing fixed effects at a sectoral level.

5.4.1 Preferential Trade Liberalization

The model results and estimation equations presented so far all referred to a unilateral trade liberalization in a simplified two-country setup[29]. While making the exposition clearer and helping to elicit the effects at work in the model, this setup is clearly not an accurate description of the reality of trade relationships in modern industrialized economies. Taking the United States as an example, while the two other countries in our data set, Canada and Mexico, are its largest trading partners, they only account for 16.6 and 13.5% of all US trade by value,

[29]Note that a bilateral trade liberalization – changing τ and τ^* by the same amount – would not lead to the discussed short- and long run changes in cost cutoffs for two countries, but instead to a decline in the cost cutoffs in both countries both in the short and the long run.

respectively[30]. Even the largest 30 US trading partners only account for about 86% of US trade, highlighting the fragmented nature of international trade. While the Melitz-Ottaviano model can be extended to an arbitrary number of countries, it is clearly not feasible to assemble a data set on all trade partners of the NAFTA countries. We do however want to recognize the multi-country structure of NAFTA by taking into account third country effects of trade barriers. Here, NAFTA can be interpreted as a preferential liberalization of Mexico *vis-a-vis* the US and Canada, as Mexico had the highest tariff barriers to start off with. In the three country case, we expect the country with the lowest sum of bilateral trade barriers to have the lowest cost cutoff, as it becomes the best export hub. To account for this, we amend equations 5.11, and 5.12 by including the relevant third country tariffs. The estimation equation for the effects of trade barriers on prices then becomes:

$$\begin{aligned}\Delta \ln\left(\frac{\bar{p}_{it}}{\bar{p}^*_{it}}\right) =& \beta_0 + \beta_1 \Delta \ln \tau_{it} + \beta_2 \Delta \ln \tau^*_{it} + \beta_3 \Delta \ln \tau^{ht}_{it} + \beta_4 \Delta \ln \tau^{th}_{it} + \beta_5 \Delta \ln \tau^{ft}_{it} + \beta_6 \Delta \ln \tau^{th}_{it} \\ & + \beta_7 \Delta \ln D_{it} + \beta_8 \Delta \ln D^*_{it} + \beta_9 \Delta \ln D^t_{it} \\ & + \gamma \left[\ln\left(\frac{\bar{p}_{it-1}}{\bar{p}^*_{it-1}}\right) + \delta_0 + \delta_1 \ln L_{t-1} + \delta_2 \ln L^*_{t-1} + \delta_3 \ln L^t_{t-1} + \delta_4 \ln \tau_{i,t-1} \right. \\ & \left. + \ \delta_5 \ln \tau^*_{i,t-1} + \delta_6 \Delta \ln \tau^{ht}_{it} + \delta_7 \Delta \ln \tau^{th}_{it} + \delta_8 \Delta \ln \tau^{ft}_{it} + \delta_9 \Delta \ln \tau^{th}_{it} \right] + \varepsilon_{ijt}\end{aligned} \tag{5.13}$$

where now h is the domestic economy, f is the foreign economy, and, with a slight abuse of notation, a t superscript denotes the third country for each country pair (e.g. when estimating the Canada-US relationship, τ^{th} are Mexican tariffs on Canadian goods, while τ^{ht} are Canadian tariffs on Mexican goods). The equation

[30] See Top U.S. Trade Partners, U.S. International Trade Administration

for productivity will be amended accordingly.

5.4.2 Dataset

Our database covers the period 1990-2007 for the North American Free Trade Agreement (NAFTA) member countries – Canada, Mexico, and the U.S. – and 64 (4-digit) manufacturing sectors. The main explanatory variable in this analysis is the tariff imposed on foreign products. All tariff data is downloaded from the World Integrated Trade Solution (WITS), an online software package published by the World Bank in collaboration with UNCTAD, the WTO, International Trade Center, and the UN Statistical Division. WITS publishes annual trade and tariff data from two different sources: the World Bank IDB database and the UNCTAD TRAINS database. Unfortunately, neither database provides a complete time series for each country that is devoid of erratic (and unexplained) jumps in the data. Thus, we created a data set that uses mostly TRAINS preferential tariff (PRF) data, but supplements it with observations from TRAINS or WTO IDB applied tariffs (AHS) where appropriate (this choice will only make a difference where there is no trade observed between countries and hence no applied rate, but a preferential rate still exists). All tariff data is reported according to ISIC Rev. 3.1 and converted to NAICS. This leads to the following rules for each country: (1) Canada: TRAINS PRF from 1989 to 1995, WTO AHS from 1996 to 2014. (2) Mexico: TRAINS AHS from 1989 to 1994, TRAINS PRF from 1995 to 2009. (3) USA: TRAINS PRF from 1980 to 1996; WTO AHS for 1997 to 2014.

For our factory gate price data, we use the producer price index (PPI) as reported by CANSIM, the Banco de Mexico, and the U.S. Bureau of Labor Statistics, respectively. All indices are normalized to equal 100 in 2003.

Labor productivity is calculated as the ratio between real value-added and total employment, as provided by the OECD SDBS database for Canada, the *Instituto Nacional de Estadistica y Geografia* (INEGI) for Mexico, and the NBER-CES Manufacturing Industry Database for the U.S. (Becker, Gray, and Marvakov, 2013). All value-added data is converted into constant 2005 USD. The number of establishments in each sector is taken from the OECD and and INEGI for Canada and Mexico, respectively, and from the Bureau of Labor Statistics for the U.S. Market size is measured by the value of GDP for each country, which is available in constant 2005 USD from the the World Bank's *World Development Indicators*. And finally, our wage data comes from the OECD SDBS database for Canada, the *Instituto Nacional de Estadistica y Geografia* for Mexico, and the NBER-CES Manufacturing Industry Database for the U.S. (Becker, Gray, and Marvakov, 2013). The Canadian and Mexican data are converted to NAICS using appropriate correspondence tables, and all values are converted to constant 2005 USD.

As discussed in the previous section, for all of the log-linearised equations, we replace the number of firms serving the domestic market, N, with the number of domestic firms producing for the domestic market, D. Unfortunately, this data is not available for all three countries during the specified time period, and thus we utilize the number of establishments, which will always be higher than the firm count as each firm may have multiple establishments. As long as the average number of establishments per firm remains constant, this should not present a problem, as our model is estimated in first differences. It is however not obvious that this relationship will remain constant in response to a trade liberalization. In fact, the main channel through which welfare gains arise in the model is the reallocation of production from unproductive to more productive firms, with less productive firms exiting the market and more productive firms expanding. If this displacement happens through larger firms taking over establishments of less productive ones, we would expect the number of establishments to stay constant, while the number of domestic producers falls (i.e. the number of establishments per firm increases). If on the other hand larger firms are simply able to expand production in existing establishments, this effect would be absent.

5.4.3 Estimation

As outlined at the beginning of Section 5.4, we follow the estimation strategy of Chen et al. (2009); however, while they use changes in domestic and foreign import penetration in sector i at time t as the main explanatory variables for changes in prices, and labour productivity, we use this as a control variable and instead rely on the domestic tariff rate (τ) imposed on foreign goods imported from the trading partner and the foreign tariff rate (τ^*) imposed on domestic goods exported to the trading partner as the main explanatory variables. To test the competitive effects of trade liberalization, we use the difference in differences approach with fixed–effects on the country–pair, industry, and year. In the short run we use the log first-difference in the explanatory and dependent variables, whereas we use a lag operator on the explanatory variables and an error correction term to estimate the dynamics in the long run.

Table 5.2 outlines the comparative statics for the theoretical model, with subscript sr denoting the ”short run” and lr denoting the ”long run”. Notice that in the long run theory suggests that the pro–competitive effects are reversed and actually take an anti–competitive nature as firms are able to relocate to new markets. Interestingly, as we will exhibit in the following section, our analysis does not provide the same long-run dynamics.

Table 5.2: Comparative Statics – Model Predictions

Regressor	**Dependent Variables**					
	$\bar{p}_{sr}$	$\bar{p}_{lr}$	μ_{sr}	μ_{lr}	z_{sr}	z_{lr}
τ	+	–	+	–	–	+
τ^*	–	+	–	+	+	–
D	–		–		+	
D^*	+		+		–	
L		–		–		+
L^*		+		+		–

5.5 Results & Discussion

Tables 5.5, and 5.7 present our results on the short-run effects of trade liberalization on prices and productivity, respectively. Column (1) in each table presents the results from our theoretical estimations in equations 5.11 and 5.12, respectively[31], with fixed effects at the country-industry pair. Column (2) shows the same regression employing fixed effects on the country-pair level, while columns (3) and (4) give results for the sub sample of free entry and fixed entry

[31] All codes used to obtain the results of this chapter are available from my GitHub repository TradeProductivity.

firms, respectively, with each regression again estimated using fixed effects at the country-industry level. Columns (5) and (6) show the regression results for the same regressions as in columns (3) and (4), albeit with fixed effects at the country-pair level.

The table shows that the effects of tariff barriers on outcomes is very imprecisely measured. While the signs on the coefficients point overwhelmingly in the direction implied by theory – higher domestic tariffs increase domestic prices in the short run, while higher foreign tariffs lower them –, only two model specifications show significant coefficients. Furthermore, the effect of domestic tariffs is more precisely estimated than that of foreign tariffs, a pattern that repeats itself when looking at the estimated coefficients on our measure of firms. Here, the foreign number of firms is generally estimated more precisely, and again all signs on the coefficients confirm the theoretical predictions. Table 5.7 repeats the analysis with relative productivity as the dependent variable. Here the results are more precisely estimated and more consistent across specifications; a higher domestic tariff lowers domestic productivity, while a higher number of domestic firms serves to increase it. Again, there is a marked difference in the precision of estimates for foreign and domestic variables, which can most likely be explained by the structure of our data set. Given that we are employing three country pairs in the estimation, not all countries contribute equally to domestic and foreign variables. Specifically, the ordering of our data set implies that Canada serves as the domestic market for two thirds of the data set (country pairs Canada-Mexico, Canada-USA), while the USA are the foreign market for two thirds of the data set. As our data set is assembled from a variety of sources, there are different patterns of missing data, and possibly different degrees of measurement error in the data from different countries. Hence, different independent variables might be affected differently by problems with the data for a specific country.

Tables 5.6 and 5.8 show the results for the inclusion of third country variables. This specification again lowers the precision of estimated effects of domestic and foreign tariffs, rendering all coefficients insignificant. There is some evidence of the effect of third country tariffs on relative prices in two countries, with domestic tariffs on third country products increasing the domestic inflation rate, while third country tariffs on foreign country imports lower the relative growth rate of the domestic price level. This provides some support for the mechanism put forward in Melitz and Ottaviano (2008), with those countries with the lowest overall sum of tariff barriers having the lowest price level. The results for the productivity equations are very similar.

Turning to the long run of the model, 5.9 and 5.11 show the results of estimating the error correction model specifications on prices and productivity, respectively. First we note that the coefficient on the lagged dependent variable is between zero and one and highly significant, indicating that the error correction specification is correct. There is little evidence of meaningful effects of trade barriers in the long run, with the exception of a negative effect of foreign tariffs on the long run price level in the domestic country, a finding that runs directly contrary to the theory. Furthermore, the effect of market size is directly opposed to what the theory predicts, with large and significant positive effects of domestic market size on domestic inflation (and large negative effects on the level of productivity). 5.11 however can be seen to vindicate the different theoretical predictions for short- and long run when viewed through the lens of market entry; columns (3), (4), (5) and (6) show that the effects of lagged variables are much stronger and the speed of mean reversion is much faster for those industries that are closer to the free entry ideal, as the model would predict.

5.6 Discussion

The only empirical application of the Melitz–Ottaviano (2008) model to date suggested that the long–run effects of trade liberalization are anti–competitive, that is, there will be a reversal in any competitive gains as firms are allowed to move to new markets. This chapter added to the evidence on the model's prediction by estimating the relationship between trade barriers, number of firms, market size and prices as well as productivity. While the model's predictions are largely confirmed in the short run, the estimates for the long run behaviour of aggregate variables provide little support for the reversion of the effect of trade barriers in the long run. Evidence in support of the theory presented here included the role of third country trade barriers in shaping relative performance between two trade partners, as well as a finding of stronger reactions to long run changes in trade barriers by industries that have high turnover rates, therefore being close to the model concept of free entry.

5.7 Appendix A: Figures, Summary Statistics, Results

Figure 5.1: Canadian Tariff on Mexican Goods

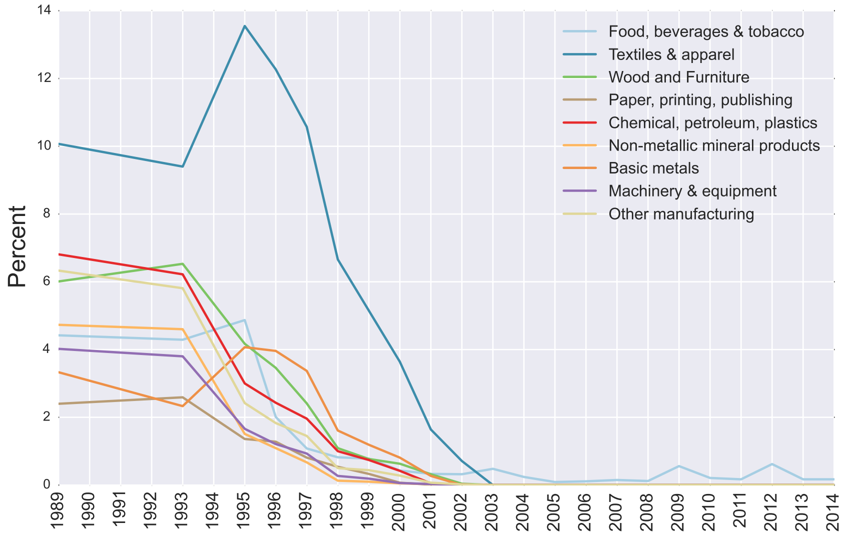
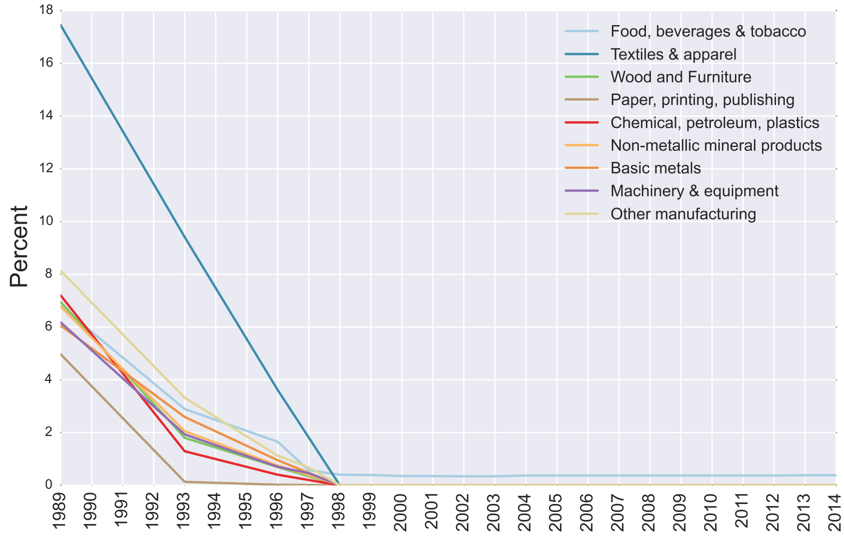


Figure 5.2: Canadian Tariff on U.S. Goods



32

[32]The majority of output tables in this paper has been produced using the stargazer package (Hlavac, 2004)

Figure 5.3: Mexican Tariff on Canadian Goods

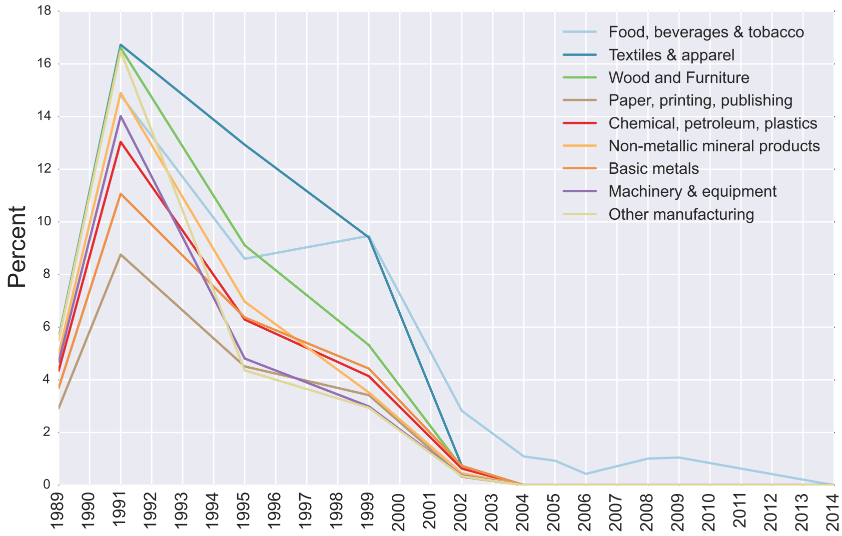


Figure 5.4: Mexican Tariff on U.S. Goods

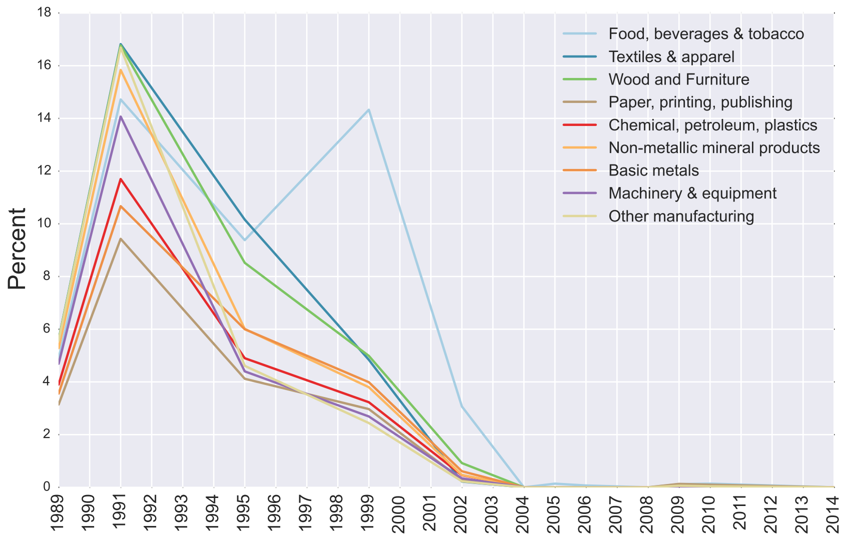


Figure 5.5: U.S. Tariff on Canadian Goods

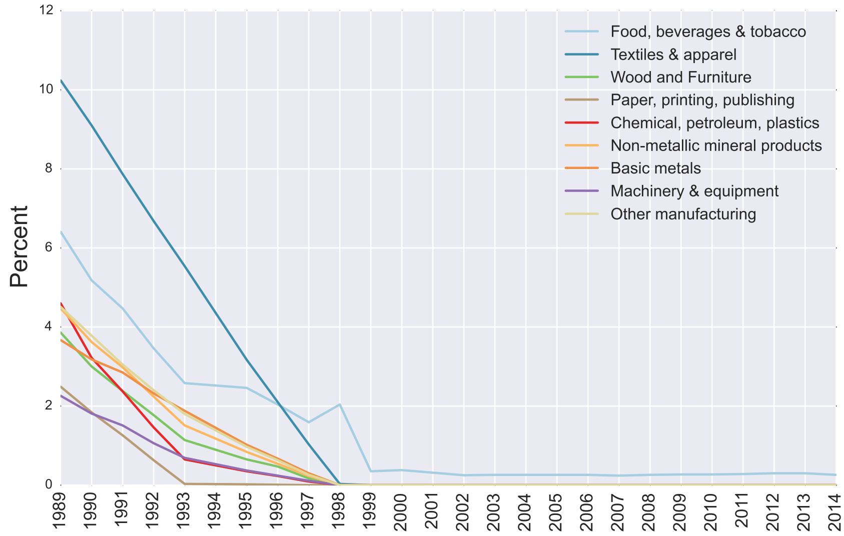


Figure 5.6: U.S. Tariff on Mexican Goods

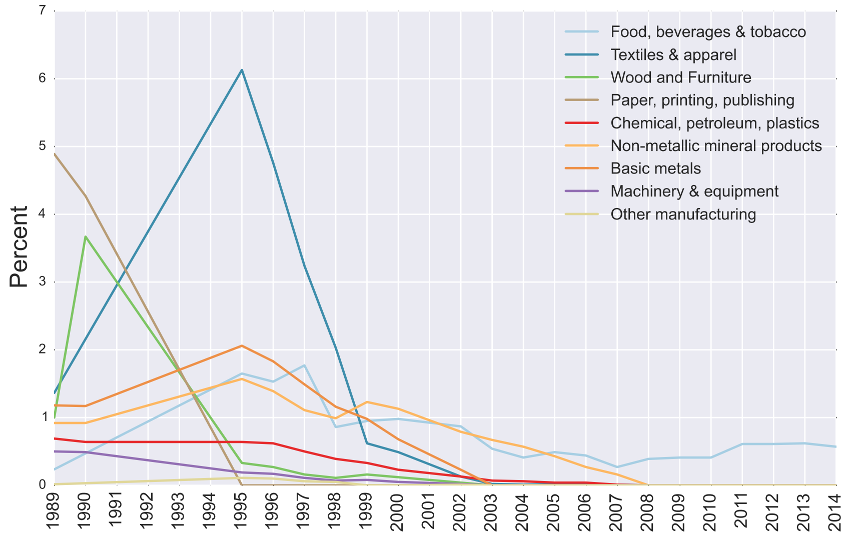


Table 5.3: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
exp_f	702	76,233,571,691.000	198,643,966,085.000	399,038,863	1,890,000,000,000
gdp_can	594	908.568	278.187	536.500	1,370.640
gdp_mex	594	1,036.795	315.456	560.660	1,566.310
gdp_usa	594	9,807.745	2,991.049	5,482.130	14,498.930
cpi_can	675	79.697	12.487	56.340	100.000
cpi_mex	675	48.219	32.559	1.560	100.000
cpi_usa	675	75.546	14.894	50.300	100.000
open_ind_can	810	59.378	12.446	37.550	75.580
open_ind_mex	810	34.822	17.307	13.210	62.320
open_ind_usa	810	22.673	3.691	17.190	30.970
ppi_mex	810	57.485	51.438	0.100	276.590
tau_s_can_mex	540	1.829	3.193	0.000	17.780
tau_s_can_usa	540	2.440	5.051	0.000	26.440
tau_s_mex_can	405	8.654	8.751	0.000	30.530
tau_s_mex_usa	405	7.611	8.552	0.000	28.800
tau_s_usa_can	621	0.994	1.780	0.000	10.600
tau_s_usa_mex	621	1.717	2.654	0.000	11.810

5.8 Appendix B: Industry list, NAICS classification

Table 5.4: Industry List, NAICS 4-digit

Industry	NAICS 4-digit
3111	Animal Food Manufacturing
3112	Grain and Oilseed Milling
3113	Sugar and Confectionery Product Manufacturing
3114	Fruit and Vegetable Preserving and Specialty Food Manufacturing
3115	Dairy Product Manufacturing
3116	Animal Slaughtering and Processing
3118	Bakeries and Tortilla Manufacturing
3119	Other Food Manufacturing
3121	Beverage Manufacturing
3132	Fabric Mills
3133	Textile and Fabric Finishing and Fabric Coating Mills
3149	Other Textile Product Mills
3152	Cut and Sew Apparel Manufacturing

3159	Apparel Accessories and Other Apparel Manufacturing
3169	Other Leather and Allied Product Manufacturing
3211	Sawmills and Wood Preservation
3212	Veneer, Plywood, and Engineered Wood Product Manufacturing
3219	Other Wood Product Manufacturing
3221	Pulp, Paper, and Paperboard Mills
3222	Converted Paper Product Manufacturing
3231	Printing and Related Support Activities
3241	Petroleum and Coal Products Manufacturing
3251	Basic Chemical Manufacturing
3254	Pharmaceutical and Medicine Manufacturing
3255	Paint, Coating, and Adhesive Manufacturing
3256	Soap, Cleaning Compound, and Toilet Preparation Manufacturing
3259	Other Chemical Product and Preparation Manufacturing
3262	Rubber Product Manufacturing
3271	Clay Product and Refractory Manufacturing

3273	Cement and Concrete Product Manufacturing
3274	Lime and Gypsum Product Manufacturing
3279	Other Nonmetallic Mineral Product Manufacturing
3312	Steel Product Manufacturing from Purchased Steel
3313	Alumina and Aluminum Production and Processing
3314	Nonferrous Metal (except Aluminum) Production and Processing
3315	Foundries
3321	Forging and Stamping
3322	Cutlery and Handtool Manufacturing
3323	Architectural and Structural Metals Manufacturing
3324	Boiler, Tank, and Shipping Container Manufacturing
3325	Hardware Manufacturing
3326	Spring and Wire Product Manufacturing
3327	Machine Shops; Turned Product; and Screw, Nut, and Bolt Manufacturing
3329	Other Fabricated Metal Product Manufacturing
3331	Agriculture, Construction, and Mining Machinery Manufacturing

3332	Industrial Machinery Manufacturing
3333	Commercial and Service Industry Machinery Manufacturing
3334	Ventilation, Heating, Air-Conditioning, & Commercial Refrig. Eq. Manuf.
3335	Metalworking Machinery Manufacturing
3336	Engine, Turbine, and Power Transmission Equipment Manufacturing
3339	Other General Purpose Machinery Manufacturing
3342	Communications Equipment Manufacturing
3344	Semiconductor and Other Electronic Component Manufacturing
3345	Navigational, Measuring, Electromedical, and Control Instrum. Manuf.
3351	Electric Lighting Equipment Manufacturing
3352	Household Appliance Manufacturing
3353	Electrical Equipment Manufacturing
3359	Other Electrical Equipment and Component Manufacturing
3362	Motor Vehicle Body and Trailer Manufacturing
3363	Motor Vehicle Parts Manufacturing
3371	Household and Institutional Furniture and Kitchen Cabinet Manufacturing

Table 5.5: Prices (Short Run), all country pairs

	Dependent variable:					
	$\Delta \log \left(\frac{p}{p^*}\right)$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \log \tau_t$	0.002	0.002	0.0002	0.004*	0.0001	0.004**
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
$\Delta \log \tau_t^*$	-0.001	-0.001	-0.001	0.0005	-0.001	0.0004
	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)
$\Delta \log D_t$	-0.029	-0.027	-0.070^{**}	0.032	-0.067^{**}	0.033
	(0.024)	(0.023)	(0.033)	(0.045)	(0.032)	(0.044)
$\Delta \log D_t^*$	0.104***	0.088**	0.082*	0.347*	0.078*	0.212
	(0.039)	(0.036)	(0.042)	(0.185)	(0.041)	(0.153)
Observations	2,769	2,769	1,021	881	1,021	881
R^2	0.004	0.003	0.009	0.010	0.008	0.008

Note: *p<0.1; **p<0.05; ***p<0.01

(1),(3): Fixed effects country-industry; (2),(4): Fixed effects country pair

3372	Office Furniture (including Fixtures) Manufacturing
3379	Other Furniture Related Product Manufacturing
3399	Other Miscellaneous Manufacturing

5.9 Appendix C: Regression Results

5.10 Appendix D: Data Appendix

Table 5.6: Prices (Short Run), all country pairs, third country variables included

	Dependent variable:					
	$\Delta \log \left(\frac{p}{p^*}\right)$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \log \tau^{hf}$	-0.001	-0.001	-0.003	0.002	-0.003	0.002
	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.002)
$\Delta \log \tau^{fh}$	0.0002	0.0002	-0.0005	0.002	-0.001	0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
$\Delta \log \tau^{ht}$	0.005^{***}	0.005^{***}	0.005	0.007^{**}	0.005^{*}	0.007^{***}
	(0.002)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
$\Delta \log \tau^{th}$	0.002	0.002	0.004	-0.002	0.004	-0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
$\Delta \log \tau^{ft}$	-0.001	-0.001	0.00002	-0.002	0.00005	-0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
$\Delta \log \tau^{tf}$	-0.007^{***}	-0.007^{***}	-0.008^{***}	-0.008^{***}	-0.008^{***}	-0.008^{***}
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)
$\Delta \log D_t$	-0.013	-0.013	-0.067^{*}	0.074	-0.062^{*}	0.072
	(0.026)	(0.025)	(0.035)	(0.048)	(0.034)	(0.046)
$\Delta \log D_t^*$	0.093^{**}	0.083^{**}	0.071^{*}	0.404^{**}	0.069^{*}	0.295^{*}
	(0.039)	(0.037)	(0.043)	(0.205)	(0.041)	(0.167)
Observations	2,522	2,522	965	743	965	743
R^2	0.023	0.022	0.028	0.044	0.027	0.040

Note: *p<0.1; **p<0.05; ***p<0.01

(1),(3): Fixed effects country-industry: (2),(4): Fixed effects country pair

Table 5.7: Productivity (Short Run), all country pairs

	Dependent variable:					
	Δ log $\left(\frac{z}{z^*}\right)$					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log τ_t	-0.009^{***}	-0.009^{***}	-0.018^{***}	-0.002	-0.018^{***}	-0.002
	(0.003)	(0.003)	(0.005)	(0.006)	(0.005)	(0.006)
Δ log τ_t^*	0.006^{*}	0.006^{*}	0.010^{**}	0.011^{*}	0.010^{**}	0.011^{*}
	(0.003)	(0.003)	(0.004)	(0.006)	(0.004)	(0.006)
Δ log D_t	0.146^{***}	0.145^{***}	0.200^{***}	0.056	0.201^{***}	0.051
	(0.056)	(0.054)	(0.072)	(0.118)	(0.070)	(0.114)
Δ log D_t^*	0.015	0.035	-0.097	0.541	-0.095	0.716^{*}
	(0.089)	(0.083)	(0.091)	(0.484)	(0.088)	(0.399)
Observations	2,695	2,695	990	860	990	860
R^2	0.007	0.006	0.022	0.006	0.021	0.008

Note: *p<0.1; **p<0.05; ***p<0.01

(1),(3),(4): Fixed effects country-industry; (2),(5),(6): Fixed effect country pair

Table 5.8: Productivity (Short Run), all country pairs, third country variables included

	Dependent variable:					
	$\Delta \log \left(\frac{z}{z^*}\right)$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \log \tau_t^{hf}$	-0.009^{***}	-0.009^{***}	-0.016^{**}	-0.005	-0.016^{***}	-0.004
	(0.003)	(0.003)	(0.006)	(0.006)	(0.006)	(0.006)
$\Delta \log \tau_t^{fh}$	-0.001	-0.001	0.003	0.006	0.003	0.005
	(0.004)	(0.004)	(0.007)	(0.008)	(0.007)	(0.007)
$\Delta \log \tau_t^{ht}$	0.001	0.001	-0.0002	0.002	-0.0003	0.001
	(0.003)	(0.003)	(0.006)	(0.007)	(0.006)	(0.007)
$\Delta \log \tau_t^{th}$	-0.005	-0.006	-0.002	-0.004	-0.002	-0.004
	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)	(0.007)
$\Delta \log \tau_t^{ft}$	0.009^{**}	0.009^{**}	0.010	0.007	0.010	0.007
	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)	(0.007)
$\Delta \log \tau_t^{tf}$	-0.014^{***}	-0.014^{***}	-0.019^{***}	-0.008	-0.019^{***}	-0.008
	(0.004)	(0.003)	(0.006)	(0.007)	(0.006)	(0.006)
$\Delta \log D_t$	0.176^{***}	0.166^{***}	0.218^{***}	0.063	0.223^{***}	0.035
	(0.058)	(0.056)	(0.075)	(0.123)	(0.073)	(0.117)
$\Delta \log D_t^*$	0.008	0.030	-0.104	0.525	-0.099	0.823^{*}
	(0.089)	(0.084)	(0.092)	(0.523)	(0.089)	(0.427)
Observations	2,506	2,506	955	741	955	741
R^2	0.022	0.021	0.042	0.011	0.041	0.013

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

(1),(3),(4): Fixed effects country-industry; (2),(5),(6): Fixed effect country pair

Table 5.9: Prices (Long Run), all country pairs

	Dependent variable:					
	$\Delta \log \left(\frac{p_t}{p_t^*}\right)$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \log \tau_t$	0.001	0.001	-0.001	0.005^{**}	-0.001	0.004^{**}
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
$\Delta \log \tau_t^*$	-0.001	0.00004	-0.00004	-0.002	0.0002	0.0004
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
$\Delta \log D_t$	-0.031	-0.025	-0.063^{**}	0.008	-0.061^{**}	0.023
	(0.023)	(0.021)	(0.031)	(0.041)	(0.030)	(0.040)
$\Delta \log D_t^*$	0.038	0.028	0.038	0.366^{**}	0.036	0.132
	(0.036)	(0.033)	(0.039)	(0.174)	(0.037)	(0.141)
$\log \left(\frac{p_{t-1}}{p_{t-1}^*}\right)$	-0.123^{***}	-0.118^{***}	-0.121^{***}	-0.149^{***}	-0.118^{***}	-0.137^{***}
	(0.006)	(0.005)	(0.010)	(0.010)	(0.009)	(0.009)
$\log \tau_{t-1}$	-0.001	-0.001	-0.001	-0.0003	-0.001	-0.001
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
$\log \tau_{t-1}^*$	-0.002	-0.00002	-0.0003	-0.006^{***}	0.0004	-0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
$\log L_{t-1}$	-0.570^{***}	-0.541^{***}	-0.497^{***}	-0.775^{***}	-0.480^{***}	-0.734^{***}
	(0.082)	(0.077)	(0.136)	(0.139)	(0.131)	(0.135)
$\log L_{t-1}^*$	0.466^{***}	0.458^{***}	0.400^{***}	0.611^{***}	0.396^{***}	0.616^{***}
	(0.080)	(0.076)	(0.132)	(0.138)	(0.126)	(0.133)
Observations	2,769	2,769	1,021	881	1,021	881
R^2	0.184	0.181	0.192	0.246	0.187	0.224

Note: *p<0.1; **p<0.05; ***p<0.01

Fixed effects for country pair

Table 5.10: Prices (Long Run), all country pairs, third country variables included

	Dependent variable:					
	$\Delta \log\left(\frac{p_t}{p_t^*}\right)$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \log \tau_t^{hf}$	-0.001	-0.001	-0.003	0.003	-0.004	0.002
	(0.002)	(0.001)	(0.003)	(0.002)	(0.003)	(0.002)
$\Delta \log \tau_t^{fh}$	0.001	0.00001	0.00003	0.0005	0.0002	0.0003
	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)
$\Delta \log \tau_t^{ht}$	0.003^{**}	0.004^{***}	0.003	0.005^{*}	0.003	0.007^{***}
	(0.002)	(0.001)	(0.003)	(0.003)	(0.003)	(0.002)
$\Delta \log \tau_t^{th}$	0.001	0.0002	0.002	0.001	0.001	-0.001
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)
$\Delta \log \tau_t^{ft}$	-0.001	-0.0002	0.001	-0.004	0.001	-0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)
$\Delta \log \tau_t^{tf}$	-0.003^{**}	-0.004^{***}	-0.004	-0.004	-0.005	-0.005^{**}
	(0.002)	(0.001)	(0.003)	(0.002)	(0.003)	(0.002)
$\Delta \log D_t$	-0.013	-0.011	-0.054^{*}	0.032	-0.050	0.046
	(0.024)	(0.023)	(0.032)	(0.041)	(0.031)	(0.040)
$\Delta \log D_t^*$	0.028	0.019	0.032	0.415^{**}	0.034	0.093
	(0.036)	(0.034)	(0.039)	(0.190)	(0.038)	(0.151)
$\log\left(\frac{p_{t-1}}{p_{t-1}^*}\right)$	-0.135^{***}	-0.127^{***}	-0.120^{***}	-0.195^{***}	-0.117^{***}	-0.168^{***}
	(0.007)	(0.007)	(0.013)	(0.013)	(0.012)	(0.012)
$\log \tau_{t-1}^{hf}$	0.00001	-0.001	-0.001	0.002	-0.001	0.0003
	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.002)
$\log \tau_{t-1}^{fh}$	0.001	-0.0003	0.0002	-0.003	0.0003	-0.003
	(0.002)	(0.001)	(0.004)	(0.003)	(0.003)	(0.002)
$\log \tau_{t-1}^{ht}$	-0.0002	0.001	-0.0004	0.002	0.001	0.003^{*}
	(0.002)	(0.001)	(0.003)	(0.002)	(0.003)	(0.002)
$\log \tau_{t-1}^{th}$	0.002	0.001	0.002	0.005^{**}	0.002	0.001
	(0.002)	(0.001)	(0.003)	(0.002)	(0.003)	(0.002)
$\log \tau_{t-1}^{ft}$	-0.003^{**}	-0.001	-0.0002	-0.006^{***}	0.0002	-0.001
	(0.002)	(0.001)	(0.004)	(0.002)	(0.003)	(0.002)
$\log \tau_{t-1}^{tf}$	0.002	0.001	0.001	0.002	0.001	0.00005
	(0.002)	(0.001)	(0.004)	(0.002)	(0.003)	(0.002)
$\log L_{t-1}$	-0.682^{***}	-0.636^{***}	-0.435^{**}	-1.136^{***}	-0.426^{**}	-1.059^{***}
	(0.113)	(0.103)	(0.191)	(0.191)	(0.176)	(0.182)
$\log L_{t-1}^*$	0.684^{***}	0.611^{***}	0.428^{**}	1.185^{***}	0.419^{**}	1.046^{***}
	(0.109)	(0.102)	(0.182)	(0.193)	(0.171)	(0.183)
Observations	2,522	2,522	965	743	965	743
R^2	0.210	0.202	0.207	0.344	0.199	0.293
Note:						$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$ Fixed effects for country pair

Table 5.11: Productivity (Long Run), all country pairs

	Dependent variable:			
	$\Delta \log \left(\frac{z}{z^*}\right)$			
	(1)	(2)	(3)	(4)
$\Delta \log \tau_t$	0.0002	-0.001	0.004	0.003
	(0.005)	(0.004)	(0.004)	(0.004)
$\Delta \log \tau_t^*$	0.001	0.004	-0.001	0.002
	(0.005)	(0.005)	(0.005)	(0.005)
$\Delta \log \theta$			-0.245^{***}	-0.214^{***}
			(0.040)	(0.042)
$\Delta \log \theta^*$			0.574^{***}	0.451^{***}
			(0.070)	(0.131)
$\Delta \log D_t$	-0.428^{*}	0.409^{*}	-0.179	0.262
	(0.246)	(0.221)	(0.210)	(0.208)
$\Delta \log D_t^*$	0.336	-0.427	0.167	-0.175
	(0.393)	(0.358)	(0.345)	(0.353)
$\log \left(\frac{z_{t-1}}{z_{t-1}^*}\right)$	-0.145^{***}	-0.348^{***}	-0.089^{***}	-0.308^{***}
	(0.016)	(0.022)	(0.015)	(0.033)
$\log \tau_{t-1}$	-0.008	-0.009^{*}	-0.007	-0.006
	(0.006)	(0.005)	(0.005)	(0.005)
$\log \tau_{t-1}^*$	0.006	0.013^{**}	0.0004	0.006
	(0.005)	(0.005)	(0.005)	(0.005)
$\log \theta_{t-1}$			-0.012	-0.092^{**}
			(0.021)	(0.043)
$\log \theta_{t-1}^*$			-0.010	0.119
			(0.021)	(0.078)
$\log L_{t-1}$	0.569	-1.531^{***}	-0.523	-1.625^{***}
	(0.623)	(0.559)	(0.539)	(0.538)
$\log L_{t-1}^*$	-0.777	1.740^{***}	0.198	1.416^{***}
	(0.636)	(0.568)	(0.547)	(0.543)
$\log w_{t-1}$	0.160^{**}	0.181^{**}	0.068	0.216^{**}
	(0.063)	(0.090)	(0.058)	(0.100)
$\log w_{t-1}^*$	-0.230^{***}	-0.584^{***}	-0.063	-0.129
	(0.069)	(0.117)	(0.062)	(0.142)
Observations	324	324	320	320
R^2	0.290	0.543	0.510	0.613

Note: *p$<$0.1; **p$<$0.05; ***p$<$0.01
Fixed effects for country pair or industry/country pair

Table 5.12: Productivity (Long Run), all country pairs, third country variables included

	Dependent variable:					
	$\Delta \log\left(\frac{z}{z^*}\right)$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \log \tau_t^{hf}$	0.008***	0.008**	0.017***	0.004	0.016**	0.006
	(0.003)	(0.003)	(0.006)	(0.006)	(0.007)	(0.006)
$\Delta \log \tau_t^{fh}$	−0.003	0.002	−0.009	−0.006	−0.006	−0.001
	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)	(0.007)
$\Delta \log \tau_t^{ht}$	−0.001	−0.005	−0.0005	0.004	−0.008	0.004
	(0.003)	(0.004)	(0.006)	(0.007)	(0.007)	(0.007)
$\Delta \log \tau_t^{th}$	−0.0003	0.002	−0.003	−0.004	−0.006	0.006
	(0.003)	(0.004)	(0.006)	(0.007)	(0.007)	(0.007)
$\Delta \log \tau_t^{ft}$	−0.001	−0.004	0.001	−0.004	0.002	−0.009
	(0.004)	(0.004)	(0.007)	(0.006)	(0.007)	(0.007)
$\Delta \log \tau_t^{tf}$	0.011***	0.011***	0.015***	0.003	0.019***	−0.001
	(0.003)	(0.004)	(0.006)	(0.006)	(0.007)	(0.006)
$\Delta \log D_t$	−0.059	−0.097*	−0.096	−0.031	−0.129*	−0.035
	(0.048)	(0.054)	(0.060)	(0.104)	(0.070)	(0.114)
$\Delta \log D_t^*$	0.170**	0.072	0.183**	−0.460	0.168**	−1.122**
	(0.072)	(0.081)	(0.072)	(0.479)	(0.085)	(0.448)
$\log\left(\frac{z_{t-1}}{z_{t-1}^*}\right)$	−0.399***	−0.052***	−0.437***	−0.325***	−0.053***	−0.045***
	(0.016)	(0.009)	(0.025)	(0.031)	(0.016)	(0.017)
$\log \tau_{t-1}^{hf}$	0.010***	0.007**	0.018***	0.004	0.009	0.007
	(0.003)	(0.003)	(0.006)	(0.005)	(0.006)	(0.005)
$\log \tau_{t-1}^{fh}$	−0.011***	−0.004	−0.021***	−0.008	−0.008	−0.003
	(0.004)	(0.003)	(0.007)	(0.007)	(0.006)	(0.006)
$\log \tau_{t-1}^{ht}$	−0.0001	0.004	−0.004	0.008	−0.004	0.014***
	(0.003)	(0.003)	(0.006)	(0.006)	(0.006)	(0.005)
$\log \tau_{t-1}^{th}$	−0.014***	−0.006**	−0.020***	−0.014**	−0.013**	−0.001
	(0.003)	(0.003)	(0.006)	(0.006)	(0.006)	(0.005)
$\log \tau_{t-1}^{ft}$	0.009***	0.002	0.012*	0.007	0.007	−0.001
	(0.003)	(0.003)	(0.007)	(0.006)	(0.006)	(0.005)
$\log \tau_{t-1}^{tf}$	0.010***	−0.002	0.017**	−0.001	0.006	−0.014***
	(0.003)	(0.003)	(0.007)	(0.006)	(0.006)	(0.005)
$\log L_{t-1}$	3.416***	2.380***	3.536***	2.687***	3.176***	0.754
	(0.206)	(0.217)	(0.320)	(0.454)	(0.358)	(0.459)
$\log L_{t-1}^*$	−3.239***	−2.204***	−3.419***	−2.636***	−3.095***	−0.604
	(0.207)	(0.219)	(0.323)	(0.468)	(0.358)	(0.469)
$\log w_{t-1}$	0.143***	−0.046***	0.122***	0.115***	−0.080***	−0.035
	(0.019)	(0.016)	(0.030)	(0.043)	(0.026)	(0.032)
$\log w_{t-1}^*$	−0.151***	0.045***	−0.146***	−0.065	0.065***	0.056*
	(0.020)	(0.016)	(0.029)	(0.047)	(0.025)	(0.033)
Observations	2,506	2,506	955	741	955	741
R^2	0.378	0.126	0.433	0.354	0.168	0.124
Note:						$^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
						Fixed effects for country pair or industry/country pair

Fixed effects for country pair or industry/country pair

Chapter 6

Conclusion

This thesis considered the effects of heterogeneity in economic models both theoretically and empirically. It reviewed the recent literature on the macroeconomics of household consumption and savings with agents that differ in their economic situation because of only partially insurable shocks to their idiosyncratic income. After presenting the workhorse Aiyagari-Bewley-Imrohoroglu- Huggett model of consumption under uncertainty, various extensions to the model and their effects on the model's predictions are discussed. Amongst these are the introduction of additional assets such as housing, the consideration of differing rates of return either because of different financial assets or because of entrepreneurial activity, an overlapping generations structure with bequest motives and institutional factors such as asset-based means testing for public insurance program.

After noting that all models of the consumer wealth distribution rely on a parsimonious AR(1) process with a persistent and transitory shock component, the thesis moves on to consider the recent literature on estimating income processes from the variance-covariance structure of earnings residuals. While this literature has a long tradition, recent work has renewed interest in the estimation of processes in which the stochastic process for income features a deterministic trend component which varies across households. Chapter 3 adds to this literature by considering the so far longest sample of the US PSID to estimate such processes, analyse their variation over time by considering sub-periods of the full sample and for the first time estimating these processes from British data from the BHPS. It documents substantial heterogeneity in the estimates obtained for different time periods and income definitions, although a common theme in the obtained estimates is that both the persistence of the AR(1) component and the variance of its innovation are significantly lower than those estimated from processes without deterministic growth rate heterogeneity.

Building on these findings, chapter 4 then used a life-cycle model of household saving based on an heterogeneous income process in which households learn about their individual specific intercept and slope parameter over the course of their working life to simulate wealth distributions. It showed that while learning is not important for the qualitative predictions of the model regarding the shape of the wealth distribution, the key drivers in model fit are the persistence of the AR(1) process and its innovation variance, precisely those parameters that in chapter 3 were estimated to be significantly lower under a HIP specification. The basic problem in fitting the model to the data is that the HIP process assigns a large part of the variability in household earning to the dispersion in growth rates, and a lower part to the permanent shock component. As permanent growth, in contrast to persistent shocks, does not require asset accumulation for consumption smoothing, the lifetime income inequality created by inequality in deterministic growth rates, does not lead to the large inequality in wealth holdings that inequality created by a volatile and persistent shock component in income does. As recent work points towards an HIP process as a good description of the actual income risk facing households, the results of models offering a good fit to the wealth distribution based on a persistent AR(1) component with large innovation variance have to be questioned.

In chapter 5, the thesis then considers the effects of heterogeneity on the supply side of the economy by testing the predictions of a model of international trade based on firms differing in their marginal productivity levels. Based on a sample of 64 industries in the three NAFTA member countries Canada, Mexico and USA, error correction models relating changes in the growth rate of tariff barriers, firms, and market size to changes in the growth rate of relative prices, productivities and markups are being estimated on country pairs. In an extension of the previous literature, the chapter also considers the effects of market entry by constructing measures of firm turnover for each industry and analysing samples of high- and low turnover industries separately, and considers the effects of third country tariff barriers on two trade partners. It finds support for the model's prediction regarding the effect of entry condition, namely industries with free entry displaying a larger reaction to changes in trade freeness in the long run, as well as a faster speed of adjustment. The effects of third country tariffs are absent in most specifications, a result that is likely to be due to the unfortunately less than complete tariff data available for some parts of the sample.

Throughout the thesis it has become clear that heterogeneity has important effects on the predictions of economic models and is crucially important when applying these models for policy analysis. At the same time, heterogeneity can be added in a variety of ways, many of which often help to explain similar patterns in the data. Here it is important to consider the effects of different dimensions of heterogeneity simultaneously, to understand their interplay and avoid ascribing too large a role quantitatively to one specific mechanism. As computer power continues to grow, more and more complex models of differences between economic agents will become feasible to solve, making heterogeneity in economics a fruitful area for future research.

Bibliography

Abowd, John M and David Card, “Intertemporal Labor Supply and Long-term Employment Contracts,” *American Economic Review*, March 1987, *77* (1), 50–68. 42, 45

Aiyagari, S Rao, “Uninsured Idiosyncratic Risk and Aggregate Saving,” *The Quarterly Journal of Economics*, August 1994, *109* (3), 659–84. 14, 24, 28, 29, 71

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare, “New Trade Models, Same Old Gains?,” *American Economic Review*, February 2012, *102* (1), 94–130. 91

_ , _ , Dave Donaldson, and Andres Rodriguez-Clare, “The Elusive Pro-Competitive Effects of Trade,” 2012. Working Paper. 91

Attanasio, Orazio and Guglielmo Weber, “Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy,” *Journal of Economic Literature*, September 2010, *48* (3), 693–751. 20

Badel, Alejandro and Mark Huggett, “Taxing top earners: a human capital perspective,” Working Papers 2014-17, Federal Reserve Bank of St. Louis July 2014. 33

Baker, Michael, "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics*, April 1997, *15* (2), 338–75. 42, 45, 46

Baldwin, John, Timothy Dunne, and John Haltiwanger, "A Comparison of Job Creation and Job Destruction in Canada and the United States," 1994. CES Research Paper. 104

Barba, Aldo and Massimo Pivetti, "Rising household debt: Its causes and macroeconomic implications–a long-period analysis," *Cambridge Journal of Economics*, January 2009, *33* (1), 113–137. 79

Bellone, Flora, Patrick Musso, Lionel Nesta, and Frederic Warzynski, "Endogenous Markups, Firm Productivity and International Trade: : Testing SomeMicro-Level Implications of theMelitz-Ottaviano Model," Working Papers 08-20, University of Aarhus, Aarhus School of Business, Department of Economics September 2008. 96

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu, "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents," *Econometrica*, 01 2011, *79* (1), 123–157. 38

Bernanke, Ben S., "The global saving glut and the U.S. current account deficit," Speech 77, Board of Governors of the Federal Reserve System (U.S.) 2005. 30

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott, "Survival of the best fit: Exposure to low-wage countries and the (uneven) growth of U.S. manufacturing plants," *Journal of International Economics*, January 2006, *68* (1), 219–237. 96

_ , _ , and _ , "Trade costs, firms and productivity," *Journal of Monetary Economics*, July 2006, *53* (5), 917–937. 101

Bewley, Truman, "The permanent income hypothesis: A theoretical formulation," *Journal of Economic Theory*, December 1977, *16* (2), 252–292. 24

Blundell, Richard and Ian Preston, "Consumption Inequality And Income Uncertainty," *The Quarterly Journal of Economics*, May 1998, *113* (2), 603–640. 25

_ , **Luigi Pistaferri, and Ian Preston**, "Consumption Inequality and Partial Insurance," *American Economic Review*, December 2008, *98* (5), 1887–1921. 25, 26

Brown, Mark W., "Renewing Canada's Manufacturing Economy: A Regional Comparison, 1973-1996," Technical Report 11 2004. 104

Browning, Martin and Thomas F. Crossley, "The Life-Cycle Model of Consumption and Saving," *Journal of Economic Perspectives*, Summer 2001, *15* (3), 3–22. 20, 22

_ , **Mette Ejrns, and Javier Alvarez**, "Modelling Income Processes with Lots of Heterogeneity," *Review of Economic Studies*, 2010, *77* (4), 1353–1381. 44

Cagetti, Marco, "Wealth Accumulation over the Life Cycle and Precautionary Savings," *Journal of Business & Economic Statistics*, July 2003, *21* (3), 339–53. 29, 31, 32

_ **and Mariacristina DeNardi**, "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, October 2006, *114* (5), 835–870. 34

Calderon-Madrid, Angel and Alexandru Voicu, "The NAFTA Tide : Lifting the Larger and Better Boats," 2007. Working Paper. 95

Carroll, Christopher D., Jiri Slacalek, and Kiichi Tokuoka, "The Distribution of Wealth and the MPC: Implications of New European Data," *American Economic Review*, May 2014, *104* (5), 107–11. 36

Castaneda, Ana, Javier Diaz-Gimenez, and Jose-Victor Rios-Rull, "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, August 2003, *111* (4), 818–857. 29, 33

Chang, Yongsung, Jay Hong, and Marios Karabarbounis, "Life Cycle Uncertainty and Portfolio Choice Puzzles," 2013 Meeting Papers 595, Society for Economic Dynamics 2013. 56

Chen, Natalie, Jean Imbs, and Andrew Scott, "The dynamics of trade and competition," *Journal of International Economics*, February 2009, *77* (1), 50–62. 16, 92, 94, 96, 99, 101, 102, 106

Chor, Davin and Kalina Manova, "Off the cliff and back? Credit conditions and international trade during the global financial crisis," *Journal of International Economics*, 2012, *87* (1), 117–133. 99, 101

Corcos, Gregory, Massimo Del Gatto, Giordano Mion, and Gianmarco I P Ottaviano, "Productivity and Firm Selection: Quantifying the 'New' Gains From Trade," *Economic Journal*, 2011, *122*, 754–798. 96

Costinot, Arnaud and Andres Rodriguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in Gita Gopinath, Elhanan Helpman, and Ken Rogoff, eds., *Handbook of International Economics, Vol. 4*, Amsterdam: Elsevier, 2014, pp. 197–261. 92

Cozzi, Marco, "Risk Aversion Heterogeneity, Risky Jobs and Wealth Inequality," Working Papers 1286, Queen's University, Department of Economics December 2014. 36

Davies, James B., Susanna Sandstrom, Anthony Shorrocks, and Edward N. Wolff, "Estimating the Level and Distribution of Global Household Wealth," Working Paper Series UNU-WIDER Research Paper, World Institute for Development Economic Research (UNU-WIDER) 2007. 30

Deaton, Angus, *Understanding Consumption* number 9780198288244. In 'OUP Catalogue.', Oxford University Press, 1992. 25

DeBacker, Jason, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos, "Rising Inequality: Transitory or Persistent? New Evidence from a Panel of U.S. Tax Returns," *Brookings Papers on Economic Activity*, 2013, *46* (1 (Spring)), 67–142. 43

DeNardi, Mariacristina, "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, 07 2004, *71*, 743–768. 36

_ , "Quantitative Models of Wealth Inequality: A Survey," NBER Working Papers 21106, National Bureau of Economic Research, Inc April 2015. 34

_ **and Fang Yang**, "Wealth Inequality, Family Background, and Estate Taxation," NBER Working Papers 21047, National Bureau of Economic Research, Inc March 2015. 40

_ **, Eric French, and John B. Jones**, "Why Do the Elderly Save? The Role of Medical Expenses," *Journal of Political Economy*, 02 2010, *118* (1), 39–75. 39

_ **,** _ **, and John Bailey Jones**, "Life Expectancy and Old Age Savings," *American Economic Review*, May 2009, *99* (2), 110–15. 39

_ **,** _ **,** _ **, and Jeremy McCauley**, "Medical Spending of the U.S. Elderly," NBER Working Papers 21270, National Bureau of Economic Research, Inc June 2015. 39

Diaz, Antonia, Josep Pijoan-Mas, and Jose-Victor Rios-Rull, "Precautionary savings and wealth distribution under habit formation preferences," *Journal of Monetary Economics*, September 2003, *50* (6), 1257–1291. 37

DixCarneiro, Rafael, "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 2014, *82* (3), 825–885. 91

Dunne, Timothy, Mark J Roberts, and Larry Samuelson, "Patterns of firm entry and exit in U.S. manufacturing industries," *RAND Journal of Economics*, 1988, *19* (4), 495–515. 104

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes, "Do the Rich Save More?," *Journal of Political Economy*, April 2004, *112* (2), 397–444. 69

Fernandez-Villaverde, Jesus and Dirk Krueger, "Consumption And Saving Over The Life Cycle: How Important Are Consumer Durables?," *Macroeconomic Dynamics*, November 2011, *15* (05), 725–770. 26, 28

Floden, Martin, "Aggregate Savings When Individual Income Varies," *Review of Economic Dynamics*, January 2008, *11* (1), 70–82. 71

Foster, Lucia, John Haltiwanger, and Namsuk Kim, "Gross Job Flows for the U.S. Manufacturing Sector: Measurement from the Longitudinal Research Database," Technical Report 2006. 104

Friedman, Milton, *A Theory of the Consumption Function*, National Bureau of Economic Research, Inc, 1957. 21

Fuhrer, Jeffrey C., "Habit Formation in Consumption and Its Implications for Monetary-Policy Models," *American Economic Review*, June 2000, *90* (3), 367–390. 37

Fukao, Kyoji, Toshihiro Okubo, and Robert M. Stern, "An econometric analysis of trade diversion under NAFTA," *The North American Journal of Economics and Finance*, March 2003, *14* (1), 3–24. 95

Gale, William G. and John Karl Scholz, "Intergenerational Transfers and the Accumulation of Wealth," *Journal of Economic Perspectives*, Fall 1994, *8* (4), 145–160. 35

Gourinchas, Pierre-Olivier and Jonathan A. Parker, "Consumption Over the Life Cycle," *Econometrica*, January 2002, *70* (1), 47–89. 29

Grossman, Gene M and Elhanan Helpman, "Trade, Innovation, and Growth," *American Economic Review*, May 1990, *80* (2), 86–91. 91

Guiso, Luigi, Michael Haliassos, and Tullio Jappelli, "Household stockholding in Europe: where do we stand and where do we go?," *Economic Policy*, 04 2003, *18* (36), 123–170. 38

Guvenen, Fatih, "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, June 2007, *97* (3), 687–712. 15, 56, 64, 67

_ , "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, January 2009, *12* (1), 58–79. 42, 43, 45, 46, 47, 50, 52

_ **and Anthony A. Smith**, "Inferring Labor Income Risk and Partial Insurance From Economic Choices," *Econometrica*, November 2014, *82*, 2085–2129. 56, 61, 62

_ **, Fatih Karahan, Serdar Ozkan, and Jae Song**, "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?," Working Papers 719, Federal Reserve Bank of Minneapolis January 2015. 43, 67, 82

Haider, Steven J, "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991," *Journal of Labor Economics*, October 2001, *19* (4), 799–836. 46

Hause, John C, "The Fine Structure of Earnings and the On-the-Job Training Hypothesis," *Econometrica*, May 1980, *48* (4), 1013–29. 42

Head, Keith and John Ries, "Rationalization effects of tariff reductions," *Journal of International Economics*, April 1999, *47* (2), 295–320. 95, 102

Hintermaier, Thomas and Winfried Koeniger, "On the Evolution of the US Consumer Wealth Distribution," *Review of Economic Dynamics*, April 2011, *14* (2), 317–338. 32, 57, 61, 62, 63, 64, 70, 71

Hlavac, Mark, "stargazer: LaTeX code and ASCII text for well-formatted regression and summary statistics tables. R package version 5.1," CRAN 2004. 120

Hoffmann, Florian, "HIP, RIP and the Robustness of Empirical Earnings Processes," 2013. Working Paper. 43

Hryshko, Dmytro, "Labor income profiles are not heterogeneous: Evidence from income growth rates," *Quantitative Economics*, 07 2012, *3* (2), 177–209. 44, 50

Hubbard, R Glenn, Jonathan Skinner, and Stephen P Zeldes, "Precautionary Saving and Social Insurance," *Journal of Political Economy*, April 1995, *103* (2), 360–99. 38

Huggett, Mark, "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of Economic Dynamics and Control*, 1993, *17* (5-6), 953–969. 14, 24

Iacoviello, Matteo, "Household Debt and Income Inequality, 1963-2003," *Journal of Money, Credit and Banking*, 08 2008, *40* (5), 929–965. 28, 79

Imrohoruglu, Ayse, "Cost of Business Cycles with Indivisibilities and Liquidity Constraints," *Journal of Political Economy*, December 1989, *97* (6), 1364–83. 14, 24

Jenkins, Stephen P., "The British Household Panel Survey and its Income Data," IZA Discussion Papers 5242, Institute for the Study of Labor (IZA) October 2010. 48

Kaplan, Greg and Giovanni L. Violante, "How Much Consumption Insurance beyond Self-Insurance?," *American Economic Journal: Macroeconomics*, October 2010, *2* (4), 53–87. 25

Kilian, Lutz, Alessandro Rebucci, and Nikola Spatafora, "Oil shocks and external balances," *Journal of International Economics*, April 2009, *77* (2), 181–194. 99, 101

Kindermann, Fabian and Dirk Krueger, "High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk," PIER Working Paper Archive 14-036, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania October 2014. 33

Kohler, Ulrich, "PSIDTOOLS: Stata module to facilitate access to Panel Study of Income Dynamics (PSID)," Statistical Software Components, Boston College Department of Economics January 2015. 47

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song, "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937," *The Quarterly Journal of Economics*, February 2010, *125* (1), 91–128. 43

Kotlikoff, Laurence J and Lawrence H Summers, "The Role of Intergenerational Transfers in Aggregate Capital Accumulation," *Journal of Political Economy*, August 1981, *89* (4), 706–32. 35

Krusell, Per, Anthony A. Smith, and Jr., "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, October 1998, *106* (5), 867–896. 36

Lawrance, Emily C, "Poverty and the Rate of Time Preference: Evidence from Panel Data," *Journal of Political Economy*, February 1991, *99* (1), 54–77. 36

Ljungqvist, Lars and Thomas J. Sargent, *Recursive Macroeconomic Theory, Third Edition*, Vol. 1 of *MIT Press Books*, The MIT Press, June 2012. 24

Lucas, Robert, "The Industrial Revolution - Past and Future," Essay, Federal Reserve Bank of Minneapolis May 2004. 13

MaCurdy, Thomas E., "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 1982, *18* (1), 83–114. 42

Mankiw, N. Gregory and Stephen P. Zeldes, "The consumption of stockholders and nonstockholders," *Journal of Financial Economics*, March 1991, *29* (1), 97–112. 38

Meghir, Costas and Luigi Pistaferri, "Income Variance Dynamics and Heterogeneity," *Econometrica*, 01 2004, *72* (1), 1–32. 44, 82

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, *71* (6), 1695–1725. 91, 96, 97

_ **and Daniel Trefler**, "Gains from Trade When Firms Matter," *Journal of Economic Perspectives*, Spring 2012, *26* (2), 91–118. 91

_ **and Gianmarco I. P. Ottaviano**, "Market Size, Trade, and Productivity," *Review of Economic Studies*, January 2008, *75* (1), 295–316. 92, 96, 97, 117

_ **and Stephen J. Redding**, "New Trade Models, New Welfare Implications," NBER Working Papers 18919, National Bureau of Economic Research, Inc March 2013. 92

Modigliani, Franco, "Life Cycle, Individual Thrift, and the Wealth of Nations," *American Economic Review*, June 1986, *76* (3), 297–313. 35

Modigliano, Franco and Richard Brumberg, "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data," in K. Kurihara, ed., *Post Keynesian Economics*, New Brunswick: Rutgers University Press, 1954. 20, 22

Moore, David W., "Half of Young People Expect to Strike It Rich," March 2003. 78

Oswald, Florian, "High Dimensional Function Approximation in Julia," `https://github.com/floswald/ApproXD.jl` 2014. 62

Ottaviano, Gianmarco I. P., Takatoshi Tabuchi, and Jacques-Francois Thisse, "Agglomeration and Trade Revisited," *International Economic Review*, 2002, *43* (2), 409–436. 97

Piketty, Thomas, "On the Long-Run Evolution of Inheritance: France 1820–2050," *The Quarterly Journal of Economics*, 2011, *126* (3), 1071–1131. 35

_ , *Capital in the 21st Century*, Harvard University Press, 2014. 14, 17, 29

_ **and Gabriel Zucman**, "Wealth and inheritance in the long run," in Anthony B. Atkinson and Francois Bourguignon, eds., *Handbook of Income Distribution*, Amsterdam: North-Holland, 2015, pp. 1303–1368. 35

_ **, Gilles Postel-Vinay, and Jean-Laurent Rosenthal**, "Inherited vs self-made wealth: Theory & evidence from a rentier society (Paris 18721927)," *Explorations in Economic History*, 2014, *51* (C), 21–40. 35

Quadrini, Vincenzo, "Entrepreneurship, Saving and Social Mobility," *Review of Economic Dynamics*, January 2000, *3* (1), 1–40. 34

Rajan, Raghuram, *Fault Lines: How Hidden Fractures Still Threaten the World Economy*, Princeton University Press, 2011. 79

Romalis, John, "NAFTA's and CUSFTA's Impact on International Trade," *Review of Economics and Statistics*, 2007, *89* (3), 416–435. 95

Saez, Emmanuel and Gabriel Zucman, "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," NBER Working Paper October 2014. 79

Siegel, Jeremy J. and Richard H. Thaler, "Anomalies: The Equity Premium Puzzle," *Journal of Economic Perspectives*, Winter 1997, *11* (1), 191–200. 37

Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron, "Consumption and risk sharing over the life cycle," *Journal of Monetary Economics*, April 2004, *51* (3), 609–633. 61

Townsend, A., "FastGaussQuadrature.jl," `https://github.com/ajt60gaibb/FastGaussQuadrature.jl` (GitHub repository) 2015. 62

Trefler, Daniel, "The Long and Short of the Canada-U.S. Free Trade Agreement," *American Economic Review*, 2004, *94* (4), 870–895. 95, 100

Vandendriessche, Damien, "Stata command to merge BHPS and Understanding Society (UKHLS)," `http://www.parisschoolofeconomics.eu/docs/vandendriessche-damien/bhps_ukhls.zip` 2015. 48

Vermeulen, Philip, "How fat is the top tail of the wealth distribution?," Working Paper Series 1692, European Central Bank July 2014. 18

Weiss, Yoram and Lee A Lillard, "Experience, Vintage, and Time Effects in the Growth of Earnings: American Scientists, 1960-1970," *Journal of Political Economy*, June 1978, *86* (3), 427–47. 42

Yang, Fang, "Consumption over the Life Cycle: How Different is Housing?," *Review of Economic Dynamics*, July 2009, *12* (3), 423–443. 27

Zeldes, Stephen P, "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," *The Quarterly Journal of Economics*, May 1989, *104* (2), 275–98. 14