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 had to lift my spirits more than once over the last four years;

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.

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.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Details of collaboration and publications:
Parts of Chapter 4 were undertaken as joint work with Ryan Weldzius.

Signature: Nils-Holger Gudat
Date: 28th of July, 2014

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 explores the implications of learning about idiosyncratic income risk on the wealth distribution and compares the model results to observed data, with a focus on the effects of changes in cross-sectional income inequality. To this end, income processes with profile heterogeneity are estimated from survey data and then used as inputs for a structural model of household saving, in which households are imperfectly informed about the stochastic process governing the evolution of their lifetime income, but can learn about the underlying parameters. Model results for a standard model are compared to those of a model with consumption habits, while a structural break in the cross-sectional variance of idiosyncratic income growth rates is employed in an attempt to capture the secular rise in income

inequality observed since the late 1970s and explore its implications for the predicted wealth distribution. 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. Our approach allows us to test the effects of NAFTA on productivity in nine manufacturing sectors in North America and thereby complement and extend the existing literature on the effects of trade liberalisations.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Distributional questions are increasingly making a comeback in economics.

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 stylized facts on the cross sectional wealth distribution in a number of countries and their evolution over time. Then, an overview of existing modeling approaches is presented,

Chapter 3 builds on the discussion in Chapter 2 by constructing a model of learning about idiosyncratic income processes and using it to simulate wealth distributions. As a first step, income processes with profile heterogeneity are estimated for different sub-periods of PSID data from 1968 to 1999 and from BHPS data in order to assess the stability of the cross-sectional variance of income growth rates across time and labor markets. The estimates are then used as inputs in a structural model of household saving first employed by Guvenen (2007), with the addition of both structural breaks in the cross-sectional variance of growth rates and habits in consumption. The models predictions for the evolution of the US consumer wealth distribution are then benchmarked against data from the Survey of Consumer Finances.

Chapter 4 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 nine manufacturing industries in the Canada, Mexico and the United States over the period of 1994 to 2010. In an extension of the approach of Chen et al. (2009), we also conduct a sub-sample analysis in which we split the same 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

Models of the wealth distribution

2.1 Introduction

Monetary policy, forecasting and derivative pricing are a handful of the many reasons that have sparked an interest in bond yields. Most modern economies utilize the term structure of interest rates to conduct monetary policy. Particularly, the short end of the yield curve is exploited to drive changes in the medium and long end of the curve. Focus is drawn to medium and long term yields due to their inherent association with borrowing costs and consequently their tight link to the economy's aggregate demand. Current yield curves bear informational content on future curves and economic activity, rendering them a potent tool for forecasting. Additionally, the valuation of complex financial instruments is often determined through interest rate models. However, despite the fact that bond prices are typically observed, bond yields need to be extrapolated by these bond prices and as a consequence, the estimation of term structure models of interest rates has spawned a wide literature due to its importance to policymakers, academics and practitioners.

Bonds, unlike other financial assets and macroeconomic variables, enjoy the peculiarity of having many observed yields associated with different maturities, at every given point in time, thus rendering both their time series and cross-sectional properties of interest. An analysis ignoring cross-sectional restrictions is possible, when focusing on a particular segment of the yield curve. However, the incorporation of cross-sectional restrictions comes with its own benefits. First and foremost, the imposition of no-arbitrage restrictions allows the extraction of risk premia by alleviating the difficulty that usually arises, that is, the inability to disentangle risk premia from expectations. Accounting for no-arbitrage introduces an additional probability measure to the physical one, known as the risk-neutral measure. By computing the difference between those two measures, one is capable to obtain estimates of the risk premium. It is important to note that the assumption of no-arbitrage is well grounded given the highly liquid nature of bond markets. In addition, these restrictions further enhance the consistency of yields across time and maturities and improve out-of-sample forecasts by reducing the number of parameters to be estimated within the model.

Having addressed the importance of working on a set of yields that vary across time and maturities, multivariate models are perceived as the appealing paradigm to capture yield dynamics. A natural response is to consider an unrestricted vector autoregression model. However, the latter is paired with the disadvantage of losing degrees of freedom due to the high-dimensionality of the model. At this point, the advantages of cross-sectional restrictions enter into play by allowing a low-dimensional factor structure to approximate the high-dimensional system. A factor structure appears to be sufficient to be able to replicate all possible shapes of the yield curve. Specifically, yield curves take different forms across time, from U-shaped curves, all the way to flat, upward or downward sloping curves. Nonetheless, typical stylized facts of yield curves include the notion that yields ought to increase with maturity, thus rendering upward sloping curves more characteristic. This fact enhances the liquidity preference theory, which stipulates that a time-varying term premium is required on long term yields to compensate for their relative lack of liquidity. Yields are also known to be highly persistent, as indicated by their strong autocorrelations. An additional trait of the yield curve is the fact that its short end is typically more volatile than its long end. This last stylized fact becomes of particular interest in today's economy, with unconventional monetary policy strategies driving short yields near their zero lower bound. By anchoring the short end of the curve, the volatility has been seen to pick up in the long end of the curve and inversely decrease in the short end. These very stylized facts aid in imposing the restrictions necessary to achieve the factor structure.

Reaching a consensus that a low-dimensional factor structure has the ability to summarize a complex and high-dimensional structure, the econometrician is now faced with a wide choice of factor structures. At this stage, it is important to note that it is widely accepted, in the literature, that three factors are typically considered sufficient (see **?**, **?**). The choice of factor structures can be synthesized in the following list of alternative models: principal components, interpolation methods and term structure models. In this chapter, arguments are made in support of the latter alternative, as it not only encompasses consistency of yield dynamics through the imposition of no-arbitrage, but it further allows the dissociation of risk premia from expectations' estimates. This chapter, thus, resumes in clarifying the ties yield curve models may have with financial and economic variables, including exchange rates, inflation and growth. This segment builds the necessary grounds for the following chapters, which apply affine term structure models of interest rates to monetary finance applications, with the aim to extract risk premia estimates.

This chapter benefits from the work of **?** and **?**, and is constructed as follows. In the second section, the basic concepts surrounding bond yields and prices are tackled. In the third section, affine term structure models are introduced. The fourth section includes a brief account of the recent developments within this literature, known as macro-finance models. The fifth section provides conclusive remarks.

2.2 Bond prices and yields

This section establishes the main definitions revolving around term structure modeling. It is important to note that term structure models focus on specific bonds, namely zero-coupon bonds. Those pay no coupons and only pay a single payoff at maturity, known as the face value of the bond, which for simplicity is assumed to amount to 1 unit of currency. Zero-coupon bonds are characterized by being purchased at discount and by the fact that they are considered as default free. Let $P_t(\tau)$ denote the price of a bond at time t that matures in τ periods and $y_t(\tau)$ denote the yield to maturity, compounded continuously. The following relationship holds.

$$P_t\left(\tau\right) = exp\left[-\tau y_t(\tau)\right] \tag{2.2.1}$$

Yields to maturity, also known as zero coupon yields, are thus naturally implied by zero coupon bond prices as follows.

$$y_t\left(\tau\right) = -\frac{logP_t\left(\tau\right)}{\tau} \tag{2.2.2}$$

Yields can also be expressed as an average of forward rates, which are the increment observed in the yield for prolonging the maturity by one additional period. The relationship of zero coupon yields and forward rates, in continuous time, is given below.

$$y_t\left(\tau\right) = \frac{1}{\tau}\int_0^\tau f_u du \tag{2.2.3}$$

In addition, by combining equations 2.2.2 and 2.2.3, the forward rate curve can be extracted by using the formula below.

$$f_t(\tau) = -\frac{P_t'(\tau)}{P_t(\tau)} \tag{2.2.4}$$

where $P_t'(\tau)$ designates the first derivative of the bond price $P_t(\tau)$. It is interesting to note that out of the three variables in question, $P(\tau)$, $y(\tau)$ and $f(\tau)$, only one of them suffices to derive the remaining two.

As previously mentioned, bond yields are not observed and need to be extracted by transforming observed bond prices. Many approaches have been taken across the years. One of them consists of the use of spline methods, including polynomial splines and exponential splines, to name a few. These were deemed dated due to their incapacity to ensure positive forward rates. **?** elaborate on this flaw by derive the yield curve using forward rates. This very method is typically used to obtain what are known as unsmoothed Fama-Bliss forward rates. The preponderance of central banks often use interpolation methods, such as the Nelson-Siegel or Svensson method, on those unsmoothed yields, in order to smoothen them. Factor models have become increasingly popular in the estimation of term structure models as they reduce the dimensionality of the problem whilst enabling the replication of all possible shapes of the yield curve. The most widespread factor designs in term structure modeling are broadly segregated into three families. The first factor structure stems from a principal component analysis, which by construction imposes factors to remain orthogonal whilst factor loadings are left unconstrained. A second structure involves interpolation methods that fit empirical yield curves. Unlike the previous method, factors remain unconstrained and factor loadings are the ones that inherit a particular empirical structure. The third category is known as the no-arbitrage dynamic term structure model. This method imposes restrictions on both factors and loadings. The most important trait of this structure revolves around the imposition of no-arbitrage restrictions on the factor loadings. Although this last class of models is very broad, the most noteworthy subclass is known to be affine term structure models. The next section comprises of a brief account of affine models.

2.3 Affine term structure models

The pricing of bonds necessitates an equivalent probability measure to the physical one $\mathbb{P}$, known as the risk-neutral probability measure, denoted by $\mathbb{Q}$. The very introduction of a second probability measure allows the imposition of the absence of arbitrage opportunities, which according to **?**, enhances estimation and forecasting efficiency as well as solidifies the consistency of the model. Assuming no-arbitrage, a bond, that pays a payoff $\Pi(T)$ at time T, is priced under the physical measure $\mathbb{P}$ using a pricing kernel M(t). The current price $\Pi(t)$ is thus the expectation of the discounted future cash flows, as seen below, where $E_t^{\mathbb{P}}$ denotes the expectation at time t under the physical measure.

$$\Pi(t) = E_t^{\mathbb{P}} \left[\frac{M(T)}{M(t)} \Pi(T) \right] \tag{2.3.5}$$

Assuming the kernel dynamics given in equation 2.3.6, where $\Gamma(t)$ and W(t) represent, respectively, the price of risk and a standard Brownian motion, the two measures, $\mathbb{P}$ and $\mathbb{Q}$, are linked through the Radon-Nikodym derivative given in equation 2.3.7.

$$\frac{dM(t)}{M(t)} = -r(t)dt - \Gamma(t)'dW(t) \tag{2.3.6}$$

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = exp\left[-\frac{1}{2}\int_t^T \Gamma(s)'\Gamma(s)ds - \int_t^T \Gamma(s)dWs\right] \tag{2.3.7}$$

It follows that equation 2.3.5 is transformed as shown below.

$$\Pi(t) = E_t^{\mathbb{Q}}\left[exp\left(-\int_t^T r_u du\, \Pi(T)\right)\right] \tag{2.3.8}$$

Let T denote the maturity of a zero-coupon bond that pays one unit of currency at maturity and $\tau = T - t$ designate the time to maturity. The instantaneous rate, denoted by r_t, is given by the limit of yields $y_t(\tau)$ as time t tends to T and the bond price is given as follows.

$$P_t\left(\tau\right) = E_t^{\mathbb{Q}}\left[exp\left(-\int_t^T r_u du\right)\right] \tag{2.3.9}$$

It is clearly reflected in equation 2.3.9 that there are two key components to modeling the yield curve, those being the existence of an equivalent measure $\mathbb{Q}$ to the physical measure $\mathbb{P}$ and the dynamics of the instantaneous rate r_t under $\mathbb{Q}$. In affine term structure models the dynamics of the instantaneous rate r_t under $\mathbb{Q}$ ought to be an affine function of the state variable X_t, which itself is an affine diffusion under the risk-neutral probability measure. The state dynamics follow an affine diffusion process, provided below:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW(t) \tag{2.3.10}$$

where the drift $\mu(X_t)$ and the variance-covariance matrix $\sigma(X_t)\sigma(X_t)'$ are affine in X_t. The drift of the state dynamics takes the following form, $\mu(X_t) = \kappa(\theta - X_t)$, where κ is the mean reversion matrix and θ represents the unconditional mean. As for the diffusion of the process, it takes the following form, $\sigma(X_t) = \Sigma s(X_t)$, where $s(X_t)$ is equal to the identity matrix in Gaussian affine models and is a diagonal matrix, of the form $s_{ii}(X_t) = \sqrt{s_{0,ii} + s'_{1,ii}X_t}$, in the stochastic volatility class of models. More is said on the latter models, given chapter 3 focuses on an exchange rate application of an affine term structure model with stochastic volatility.

Bond prices thus inherit an exponentially affine representation, which is the solution of a system of ordinary differential equations (ODE). These ODE have a closed-form solution when the model is Gaussian and are solved numerically when the model encompasses stochastic volatility.

It is important to note that Gaussian affine models do not preclude interest rates from being negative. This issue is not of particular interest when interest rates are at a safe distance of the zero lower bound. However, with recent economic developments, interest rates have plummeted to unprecedented levels, sparking thus the need to impose the non-negativity of interest rates. Three different classes of models have been developed to accommodate this situation: shadow rate models, Cox-Ingersoll-Ross models and quadratic models. Quadratic models as in **?** are, nonetheless, unable to conform to prolonged periods of zero or near zero interest rates. Conversely, shadow rate models are able to cope with extended periods of near zero rates by rendering instantaneous rates non-linear. Chapter 4 elaborates on the particularity of estimating rates in the vicinity of zero and builds an inflation application of both a Gaussian affine term structure model and a shadow model.

It is worth noting that affine models, despite their advantages in precluding arbitrage opportunities and obtaining known expressions for term premia, come at the disadvantage of being hard to estimate and interpret. More specifically, common issues that arise are the inability to interpret intuitively the latent factors and the global optimum problem.

2.4 Macro-finance extensions

The two previous sections have established that term structure models are of importance to model the dynamics of yields across both their cross-section and time series and are particularly interesting tools due to their simplicity in extending them to more complex and complete frameworks. It has long been instilled that the state of the economy has an impact on financial variables. A clear example of macroeconomic variables feeding financial variables is the effect of the level of inflation on the future bank rate, which eventually translates to all yields in the market. Nonetheless, it is becoming increasingly apparent that the health of financial and banking institutions can have an effect on economic variables. The advent of the recent financial crisis has thus strengthened the relation between financial and economic variables, rendering macro-finance models of great importance. This section analyzes the recent developments in the use of term structure models of interest rates to macroeconomic and financial applications.

The most natural account of a macroeconomic model is the Taylor rule, which accounts for fluctuations in short rates by using the output gap and inflation gap which are the dispersion of actual levels of output and inflation, respectively, from their target values. **?** estimate a Taylor rule and are able to draw the monetary policy shocks by imposing cross-sectional restrictions. An interesting attempt of a macro model is made by **?**, who model the yield curve using level, slope and curvature factors as well as observable macroeconomic variables, amongst which are monetary policy tools, inflation and real activity. **?** have a similar approach by analyzing the effect of the inclusion of macroeconomic variables on the forecasting of the term structure of interest rates. Reported results suggest that accounting for macroeconomic informational content improves the forecasting of yields.

On the finance end of the spectrum, **?** examine the interrelation between the expected excess returns on bonds and equity and find that changes in these expected excess returns, real yields and risk levels bear a predictable component. Similarly, **?** expand upon this idea by jointly pricing the term structure of interest rates, the risk-return levels of stocks and the returns on the aggregate market.

A recently popular extension of the term structure literature consists in shedding some light on the following twofold research questions. Does the yield curve span yields' volatility, or is volatility unspanned? Those inquisitions have been triggered by a very common phenomenon in the term structure literature, that is the inability of models to jointly capture the first and second moment of yields. **?** examine whether bonds do span the yield volatility and find arguments against this hypothesis. Their conclusion was supported by the fact that yield volatility factors were uncorrelated to the yields' cross-section. According to **?**, volatility is said to be unspanned if bonds are unable to hedge the volatility risk. On this front, it is found that current unspanned stochastic volatility models cannot capture the cross-section of bond volatility. Moreover, **?** assess whether macroeconomic content has a predictive ability on the yield curve and on excess bond returns. The use of macroeconomic variables is extended to both the obtention of yield curve factors and the identification of the sources of risk which are not hedged by bonds. Therefore, spanned and unspanned stochastic volatility is a potentially prolific strand of the term structure literature which necessitates further investigation and requires further advances in the years to come.

Interesting extensions to term structure models can be found in the two types of vector autoregression (VAR) models that follow. The first consists in studying term structure models in a global scale, in the spirit of **?** that fits the yield curve of multiple countries by featuring global and country-specific factors. Similarly, **?** introduce the Global VAR model (GVAR). This paper studies the joint forecasting of financial and macroeconomic variables at an international level. Advances in the literature are expected to be made on the selection and number of global factors and individual factors. Additional consideration ought to be made on the existence of regional factor structures. An alternative is to use a Bayesian VAR (BVAR) à la **?**. This paper,with the help of artificial data, uses a term structure model as a prior. This approach allows the loose imposition of no-arbitrage conditions whilst further alleviating the dimensionality problem and accounting for possible model misspecifications.

2.5 Conclusion

This chapter provides a brief and concise outline of term structure models, covering basic concepts and introducing several advances within this literature. The general idea that transcends within the chapter is the complexity involved in estimating the term structure of interest rates as well as their potency in extracting information regarding macroeconomic and financial variables. The two following chapters will utilize term structure models in order to extract risk premia. Specifically, chapter 3 emphasizes on the link between term structure models and currencies whilst chapter 4 concentrates on the strong relationship between the yield curve and inflation. Both chapters emphasize on the affine class of term structure models and more specifically on a Nelson-Siegel affine term structure model which further imposes no-arbitrage conditions to ensure the consistency of yield dynamics.

Chapter 3

The implications of learning about income processes in the face of structural breaks for the household wealth distribution

3.1 Introduction

What factors can explain the substantial and persistent increase in US household indebtedness over the last three decades? This question has inspired a large literature that has put forward several explanations. On the side of credit supply, the institutional framework plays a central role; Campbell and Hercowitz (2005) emphasize the effects of the liberalization of the US home mortgage market in the early 1980s while Narajabad (2012) points towards improvements in banks' risk assessment capabilities to explain an increase in credit lines available to households. On the demand side, the most influential view is arguably given by Krueger and Perri (2006) who interpret the expansion of US credit markets as the rational response of consumers borrowing in order to smooth out income shocks, the variance of which has risen over time. However, this view rests on the observed increase in cross-sectional income variation being largely due to a rise in the variance of transitory idiosyncratic shocks, an interpretation that has been called into question by several authors. As an example, Kopczuk et al. (2010) find that the rise in income inequality was almost entirely driven by increases in permanent earnings inequality, with no mitigating effect of mobility across income groups that decreased at the same time as earnings inequality increased. In the same vein, DeBacker et al. (2013) use a confidential panel of tax returns from the IRS to show that all of the rise in the variance of male labor earnings between 1987 and 2009 can be attributed to a rise in the variance of the persistent part of income. In light of these difficulties, the present work aims to examine an additional mechanism that could drive household credit demand – consumption habits. In this regard, this work bears some resemblance to the prominent Rajan (2011) hypothesis that the foundations for the financial crisis of 2007/08 were laid by a credit expansion that mostly benefited low-income households. This also relates to more heterodox, post-Keynesian and Marxist explanations of rising household indebtedness that emphasize the role of debt as a substitute for stagnating or declining real wages in the middle and lower percentiles of the income distribution (see, for example, Barba and Pivetti, 2009), a mechanism of course that would require households to be – at least for some time – oblivious towards the realities of the path of their income stream. Often, behavioral explanations such as conspicuous consumption or household optimization based on a relative income hypothesis are invoked, examples include Bertrand and Morse (2013), who find empirical support for the hypothesis that consumption of lower income households is influenced by consumption of high income households, and van Treeck (2012), who presents some calculations based on Duesenberry's (1940) relative income hypothesis. Here, we try to frame this argument in terms of an arguably more standard economic model: a life-cycle model with fully optimizing rational households that derive utility from current and past consumption, while facing an uncertain income stream that they learn about over the course of their working life. The basis for the present analysis is a model introduced by Guvenen (2007), in which the income process consists of a permanent stochastic AR(1)-component, as is standard in most of the literature, and an additional deterministic term, that is different across workers and that they have to learn about in order to make precise forecasts of their lifetime income. This uncertainty, coupled with the time-inseparability of consumption introduced by habit formation, can under certain circumstances lead to an increase in the indebtedness even of households with permanently low incomes.

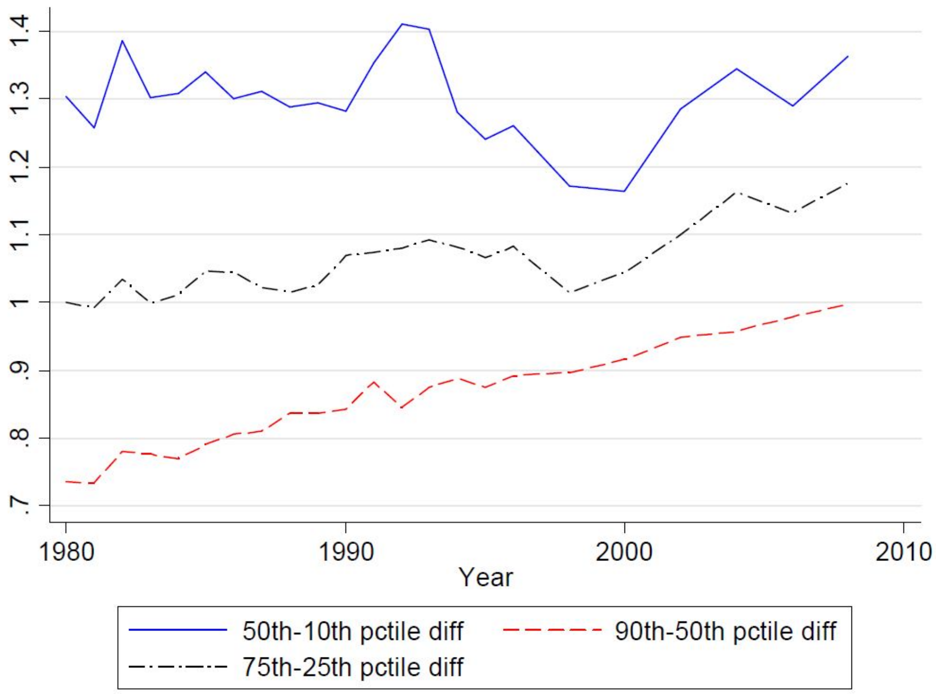
The remainder of this paper is structured as follows: Section (3.2) presents some stylized facts on the developments of the US income and wealth distribution that motivate this work. Section (3.3) lays out the model and gives analytical results for a stripped down version of the model that shed light on the mechanisms through which habit formation can generate higher indebtedness. Section (3.4) presents the results of a quantitative evaluation of the model that is informed by data from the CEX and the SCF. Section (3.5) extends the model to include a labor supply decision and gives results for different assumptions on prior uncertainty, learning behavior and the specification of the habit formation process. Section (4.5) discusses the success of different versions of the model in replicating observed data. Section (**??**) concludes.

3.2 Stylized Facts

The widening inequality in the U.S. income distribution is a well documented feature of the data during at least the last three decades. Numerous studies have examined the secular rise in top incomes (Piketty and Saez, 2003) and the flattening path of middle and lower incomes (Autor et al., 2005) and put forward explanations such as changes in relative demand and supply for different skill levels, the decline of union power, increases in international trade and competition (Ma, 2013) and the fall in the real value of the federal minimum wage[1]. While the exact timing and magnitude of the rise in inequality may differ slightly from one data source to the other and depending on the exact definition of income employed, its existence can be regarded as a consensus in the literature. Figure (3.1) is taken from Attanasio, Hurst and Pistaferri (2012) and shows the evolution of income inequality at different points of the income distribution from 1980 to 2010 based on PSID data. The rise in overall inequality is apparent and can be seen to be mostly driven by a surge in top incomes (even though, as the authors note, PSID data undersamples very rich households and thus most likely understates the rise in top incomes). There exists less consensus about the evolution of consumption inequality over the same period. While early prominent studies such as Krueger and Perri (2006) used CEX data to argue that there has been virtually no increase in consumption inequality and built theoretical models that could account for this puzzle, more recently other authors have found a larger increase in consumption inequality using different data sources that arguably suffer from less measurement error than the aggregate CEX data. Heathcote et al. (2010) use CEX spending data to document a very modest increase

[1] Some doubts on the magnitude and timing of the rising inequality are raised by Gordon (2009).

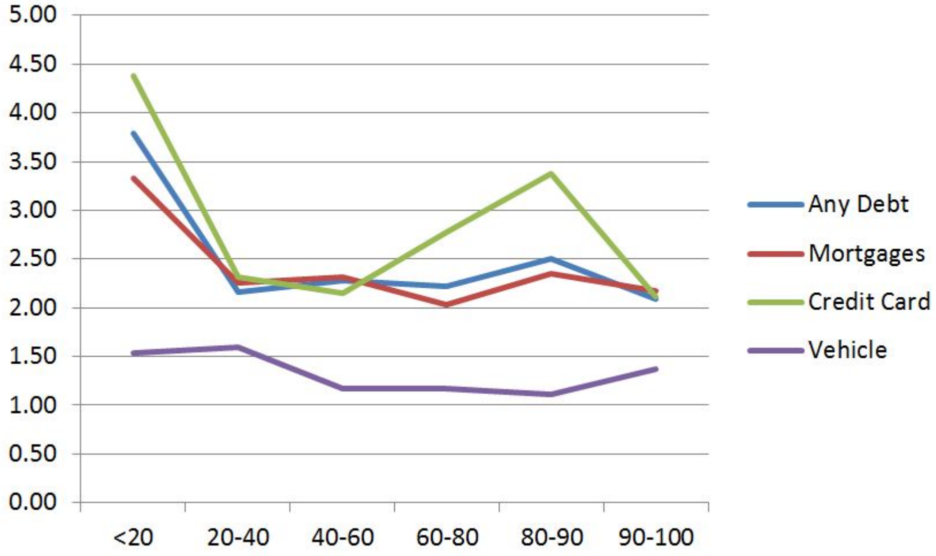
Figure 3.1: Income inequality for three income differentials, 1980-2010, PSID data



in consumption inequality, while Aguiar and Bils (2011) use the difference between income and spending in CEX data to document a rise in consumption inequality that is almost as large as the rise in income inequality and Attanasio et al. (2013) argue that, looking at some of the more precisely measured items in the CEX, one finds that consumption inequality has indeed tracked income inequality. For the purpose of this paper, we will use aggregate spending data from the CEX and thus assume that consumption inequality has not risen to the same extent as income inequality while keeping in mind that this is not a foregone conclusion. We will return to this point in the last chapter.

Another well documented feature of the data is the rise in debt holdings of the private sector. As with the rise in income inequality, this development has been widely noted and discussed with numerous explanations put forward, including changes in the regulatory framework and banking technology that widened credit supply. Most relevant to this work is the demand side argument of Krueger and Perri (2006), who explain the rise in indebtedness with a limited commitment model in which the variance of the transitory component rises and hence more insurance is required, and, indeed, optimal from a welfare perspective. However, as Cordoba (2008) points out, their model produces to empirically unappealing results: for one, wealth holdings are not concentrated at the top of the distribution, and second, the model predicts a large fraction of agents in the economy with negative wealth holdings, when their number really is close to zero in the data. Furthermore, their argument is weakened by a number of studies that find changes in the variance of the permanent component of income shocks to be the driving force behind the rise in income inequality. This is hard to reconcile with the fact that individuals at the lower end of the income distribution are those that increased their debt holdings the most. Figure (3.2) shows the changes in debt holdings for various debt categories calculated from SCF data for 1989 to 2007. While the large rise in indebtedness for the lowest income group is an artifact of business owners with failed businesses in the data, it is interesting to note that the rise in overall debt has been at least as large for the 20th to 60th percentile as for the highest income percentile. This comes as a surprise if one takes into account the higher income growth rates for individuals in higher income quintile since the early 1980s. Assuming that the dispersion in incomes is mostly driven by an increase in dispersion in the permanent component of income (as suggested by, among others, Kopczuk et al., 2010), a standard life-cycle model would suggest that while high income households should borrow against their higher future income, low income households wouldn't have an incentive to borrow when wages are stagnant. Furthermore, note that the data is not scaled by income,

Figure 3.2: Changes in debt holding by income percentile, 1989-2007, Source: SCF



so that debt holdings relative to income have increased a lot more for lower income households, given that their income growth rates over the same period have been lower than those of higher income households. Indeed, Barba and Pivetti (2009) find that installment loans and credit card debt amount to 59% of disposable income for households in the lowest income quintile of the 2004 SCF, while they amount to only 11% for the highest quintile.

One adjustment margin for households that could explain a decoupling of income and consumption inequality at least over short to medium frequencies is obviously saving. Saez and Zucman (2014) find that the share of wealth holdings of the bottom 90% of the wealth distribution has fallen from 35% in the mid 1980s to 23% in 2012, citing low growth of middle-class income, financial deregulation leading to predatory lending and behavioral biases in savings decisions as possible explanation. This paper can be seen as an attempt to investigate to what extent imperfect information of agents can account for the observed change in the wealth distribution.

- Increase in the dispersion of wealth holdings that tracked or exceeded the rise in income inequality
- Decreasing aggregate savings rate
- Prevailing misperceptions about future economic situation (Moore 2003)

3.3 The Model

The modelling approach draws heavily on the approach in Guvenen (2007) and mostly follows the notation therein. As in Guvenen's work, consumers maximize utility from consumption over the life-cycle while facing an uncertain income stream that consists of a common, known component capturing experience effects, an unknown, individual specific linear term that agents have to learn about, an AR(1) component and a purely transitory shock. The important departure in this work is that consumers also build a stock of habit that enters the utility function multiplicatively, as is standard in the literature to avoid numerical problems with CRRA utility functions (for a discussion, see Carroll (2000)). Consumers maximize

$$E_0 \left[\sum_{t=0}^{T} \beta^t \frac{\left(c_t^{1-\gamma} \left(\frac{c_t}{h_t} \right)^{\gamma} \right)^{1-\sigma}}{1-\sigma} \right] \tag{3.3.1}$$

s.t.

$$h_{t+1} = (1-\lambda)h_t + \lambda c_t \tag{3.3.2}$$

$$a_{t+1} = (1+r)a_t + y_t - c_t \tag{3.3.3}$$

$$y_t^i = g(\theta^0, X_t^i) + f(\theta^i, X_t^i) + z_t^i + \epsilon_t^i \tag{3.3.4}$$

$$a_{t+1} \geq \underline{a} \tag{3.3.5}$$

where c_t is consumption in period t, h_t is the habit stock accumulated up to period t, a_t are asset holdings subject to a borrowing constraint $\underline{a}$ – the specification of which will not be straightforward in this model and requires some further discussion

– and y_t^i is individual income, that can further be broken down into several parts:

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

where $g(\theta^0, X_t^i)$ 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 labor market, meaning that in order to calculate an expected lifetime income to base consumption (and thus, implicitly, habit) choices on, consumers in the model have to hold beliefs over the values of their individual parameters. Here again we follow Guvenen in assuming that these beliefs are updated in a Bayesian fashion, which means beliefs are updated by 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, we can derive an optimal forecast based of Bayesian updating for the evolution of the individuals belief about the evolution of

S_t^i as

$$\hat{S}_{t|t}^i = \hat{S}_{t|t-1}^i + P_{t|t-1}H_t[H_t'P_{t|t-1}H_t + R]^{-1}(y_t^i - H_t'\hat{S}_{t|t-1}^i) \tag{3.3.6}$$

$$\hat{S}_{t+1|t}^i = F\hat{S}_{t|t}^i \tag{3.3.7}$$

where $P_{t|t}$ 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 $P_{t+1|t}$:

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H_t[H_t'P_{t|t-1}H_t + R]^{-1}H_t'P_{t|t-1} \tag{3.3.8}$$

$$P_{t+1|t} = FP_{t|t}F' + Q \tag{3.3.9}$$

With Q denoting the covariance matrix of the innovation in the state space representation of $\hat{S}_{t+1|t}^i$ (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:

$$V_t(a_t, h_t, y_t^i, \hat{S}_{t|t-1}^i) = \max_{\{c_t^i, a_{t+1}^i\}} \left\{ u(c_t, h_t) + \mathbb{E}_t \left[V_{t+1}(a_{t+1}, h_{t+1}, y_{t+1}^i, \hat{S}_{t+1|t}^i | \hat{S}_{t|t-1}^i) \right] \right\} \tag{3.3.10}$$

which again has to be solved subject to the constraints 3.3.2, 3.3.3, 3.3.5 and 3.3.6-3.3.9. At the time

It can be shown that[2], for general specifications of the utility function, the Euler equation of the problem is:

$$u_t^c = \mathbb{E}_t \left[(1+r)\beta \left[u_{t+1}^c + \beta(u_{t+2}^h - (1-\lambda)u_{t+2}^c) \right] \beta \left[u_{t+1}^h - (1-\lambda)u_{t+1}^c \right] \right]$$

[2] A detailed derivation can be found in Carroll (2000).

3.3.1 Computational Algorithm

As described by Guvenen (2007), the model requires a unique construction of the state space to be solved successfully. Here we describe the procedure adapted for our purposes:

i. Draw 100 different (α^i, β^i) combinations from a bivariate normal distribution with mean $\binom{2}{0.02}$ and a variance-covariance matrix taken from PSID data. Rank the resulting types by β. For the policy experiment, increase β^i for agents above a cutoff level i^* at period T^*.

ii. For each type i, simulate 1000 income histories from $y_it = \alpha^i + \beta^i_t t + \rho z_{i,t-1} + \epsilon^i_t$

iii. Use the Kalman filter to derive, specifying some initial beliefs, a belief vector $\hat{S}^i_{t|t}$ for each agent at each t.

iv. Divide the space spanned by the derived beliefs, $[\alpha_{min}, \alpha_{max}] \times [\beta_{min}, \beta_{max}] \times [z_{min}, z_{max}]$, into 8000 equally sized cubes by taking 21 points along each dimension. At each t, if any of the $(\hat{\alpha}^i, \hat{\beta}^i, \hat{z}^i)$ points fall into one of the cubes, assign a grid point to the center of the cube and discard all empty cubes. This will result in Q_t vectors $\tilde{S}^q_t = (\hat{\alpha}, \hat{\beta}, \hat{z})^q$, $q = 1, ..., Q_t$, where Q_t is the number of non-empty cubes at t.

v. To make the income grids consistent with the bounds set by (α^i, β^i), construct one grid $y^q_{grid} = [y^q_{min}, y^q_{max}]$ per belief vector $\tilde{S}^q_t$, where the bounds are given by $\exp[H\tilde{S}^q_t \pm 3\sigma(\tilde{y}^q)]$, with σ taking from the posterior variance of the forecast of y based on the Kalman filter. Place 8 equidistant nodes on this grid.

vi. Construct a wealth grid with 12 points, more densely spaced around the borrowing constraint.

vii. Construct a habit grid with 15 points, more densely spaced around average lifetime consumption.

viii. Solve the optimization backwards from T for each t on a $[12 \times 15 \times 8 \times Q_t]$ grid.

An issue to consider when imposing the wage structure above is household's relative position in the wage distribution over the life cycle. As Kopczuk et al. (2010) and DeBacker et al. (2013) have shown, the rise in inequality was mostly persistent, and Jantti and Jenkins (2013) show that over ten year periods the largest fraction of households remains in a given income percentile. Further, a number of papers (Card et al. (2013), Bernard and Jensen (1995), Autor et al., 2008) show that a large part of the rise in income inequality can be accounted for by a rise in inequality in pay between firms and industries. Hence, even if households fully understand the changes in the wage structure favoring higher earning occupations, reacting to that change might be impossible to the extent that human capital accumulated over the working life is firm- or occupation specific [3]. For these reasons, we believe that fixing β for each household is not too restrictive.

[3] In fact, Kambourov and Manovskii (2009) show that households switching occupations are contributing to the flattening of earnings profiles at the lower end of the distribution as the transfer from one occupation to another destroys human capital and hence earnings potential.

3.4 Quantitative Results

Given the slow learning induced by the nature of the income process, the initial beliefs of consumers are of crucial importance for consumption decisions in the first periods of live. These in turn determine a habit stock that might (depending on the parameter choice for λ) have a long lasting effect on the marginal utility of consumption in the following periods. Hence, it is important to explore the sensitivity of results to different initial belief vectors and think about reasonable parametrizations. In a more fully specified model, one could further introduce a trade-off for agents between "comforting" expectations about their own future and the cost attached to acting on overly optimistic preferences, as for example argued by Glaeser (2004), but such a specification would require the introduction of a further unknown parameter in the utility function determining the utility of optimistic expectations and is therefore outside the scope of this work. Instead, we will focus on two baseline cases and explore the sensitivity of results to deviations from this baseline.

The first case assumes that agents entering the labor market in the model at age 25 have formed expectations based on previous observations of wage growth for workers in similar occupations to the one they are entering. Hence, we will assume that someone entering the labor market at, say, the 20th quintile of the wage distribution, will have expectations that were formed on the wage growth of workers in the 20th quintile 10 years prior to the worker entering the labor market. Note that this will have opposite effects on workers entering the labor market in the late 1970s, just before cross-sectional wage dispersion started to increase: while workers in low-skill occupations at the bottom of the wage distribution will face lower income growth rates than their predecessors, and thus growth rates below their expectations, workers at the high end of the wage distribution will see their incomes grow above expectations for the same reason. Thus, with this parametrization, poorer households will build up habit stocks that are too high relative to lifetime income early in life, with the reverse being true for high income households.

The second baseline case is inspired by the aforementioned research on people's inclination to hold optimistic beliefs as well as survey evidence on the overconfidence of economic actors. One such example would be a Gallup poll (Moore, 2003) in which 31% of respondents declared to expect to be rich at some point in their life, a number that jumps to 51% for the group of 18 to 29 year olds, where *rich* is defined as having an annual income of more than $120,000 or assets in excess of $1,000,000. DiPrete (2007) surveys a number of similar polls, compares their results with PSID data and concludes that even accounting for subjective differences in the definition of "being rich", Americans significantly overestimate the opportunity for upward income mobility over their lifetime [4]. A host of similar studies can be found and while one can certainly question whether such obviously unreasonable expectations form the basis for everyday economic decisions, they do point towards a significant amount of unwarranted optimism about the own economic future for a large part of the population. With this in mind, in the second scenario the belief vectors are parametrized to values that exceed the realized growth rates for all income brackets over the entire sample period. While comparing these two cases already gives us a good deal of information about the sensitivity of results to the belief vector, in section (4.5) we will also discuss the results using Guvenen's (2007) parametrization

[4] Interestingly, this overconfidence seems to have been dampened by the financial crisis, if more recent studies are an indication. Compare, e.g. http://www.cnbc.com/id/44559645

to make the model results directly comparable and of experimenting with more extreme values of beliefs.

3.5 Extensions

An important part of the mechanism explored quantitatively above was the interaction between profile heterogeneity, initial beliefs and learning, the speed of which in turn depended on the nature of the income process underlying the simulations. Obviously, estimating the model using a more standard income process without profile heterogeneity, as is done in most of the life-cycle literature, would eliminate this mechanism completely, as there would be no meaningful learning about the income process. However, recent work by Hoffmann (2013) shows that specification error in the income process can severely bias econometric results towards either over- or underestimating the importance of profile heterogeneity. Hoffmann proposes a more general and flexible specification, which is still tractable enough to be used in a dynamic programming problem.

3.6 Discussion

3.6.1 Extensions

- Health inequalities: The model does not account for the fact that there is substantial variation in life-expectancy across the population with a strong correlation between life expectancy and income/asset level. This could lead to higher aggregate savings than in a model that does not account for this fact, as higher income households have to save a larger fraction of income during their working life to insure longevity risk than low income households.

Chapter 4

The competitive effects of trade liberalisation in North America: An empirical application of the Melitz and Ottaviano Model

4.1 Introduction

The economic benefits of free trade are arguably one of the most uncontroversial result 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 international trade [5]. 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)).

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 lass class of trade models, 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

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

gains from firm selection, a fact that motivates us to test the model directly in data on international trade. 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 models equilibrium conditions that allow us to test the effects of trade openness on relative price levels, markups and labor 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 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 nine manufacturing sectors in Canada, Mexico and the US, covering the time period from the introduction of NAFTA in 1994 up to 2006, 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.[6].

[6]Given that the US already engaged in a free-trade agreement with Canada in 1988 (CUSFTA),

Our findings support the main model predictions, with tariff barriers stifling domestic competition, leading to higher producer prices, markups and 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 4.2 gives a survey of previous literature assessing the effects of trade liberalizations in general and of NAFTA specifically. Section 4.3 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 4.4 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 4.5; Section 4.6 concludes.

we also estimate the short-run effects of this agreement prior to NAFTA and the results hold

4.2 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 labor 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

(2006), 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.

4.3 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 cut-offs in two markets given by:

$$N = \bar{N}\left(\frac{c_D}{c_M}\right)^k + \bar{N}^* \frac{1}{\tau^k}\left(\frac{c_D}{c_M^*}\right)^k \tag{4.3.1}$$

$$N^* = \bar{N}^*\left(\frac{c_D^*}{c_M^*}\right)^k + \bar{N} \frac{1}{(\tau^*)^k}\left(\frac{c_D^*}{c_M}\right)^k \tag{4.3.2}$$

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

$$c_D = \left[\frac{\phi c_M^k}{L} \frac{1-\tau^{-k}}{1-(\tau\tau^*)^{-k}}\right]^{\frac{1}{k+2}} c_D^* = \left[\frac{\phi c_M^k}{L} \frac{1-\tau^{-k}}{1-(\tau\tau^*)^{-k}}\right]^{\frac{1}{k+2}}, \tag{4.3.3}$$

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:

$$\left(\frac{\bar{p}}{\bar{p}^*}\right)^k = \left(\frac{c_D}{c_D^*}\right)^k = \left(\frac{c_M}{c_M^*}\right)^k \frac{\bar{N}*}{\bar{N}} \frac{N}{N^*} \frac{1+\frac{\bar{N}}{\bar{N}^*}\frac{1}{(\tau^*)^k}\left(\frac{c_M^*}{c_M}\right)^k}{1+\frac{\bar{N}^*}{\bar{N}}\frac{1}{\tau^k}\left(\frac{c_M}{c_M^*}\right)^k} \tag{4.3.4}$$

and in the long run:

$$\left(\frac{\bar{p}}{\bar{p}^*}\right)^(k+2) = \left(\frac{c_D}{c_D^*}\right)^(k+2) = \left(\frac{c_M}{c_M^*}\right)^k \frac{L^*}{L} \frac{1-\frac{1}{\tau^k}}{1-\frac{1}{(\tau^*)^k}} \tag{4.3.5}$$

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 4.3.4, 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:

$$\frac{1}{\tau^k}\left(\frac{c_M}{c_M^*}\right)^k = \frac{\theta}{1-\theta} \tag{4.3.6}$$

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 labor productivity: $\bar{c} = \frac{w}{z}$. If it is additionally assumed that unit labor costs only depend on nominal wages, the ratio of domestic to foreign labor productivity can be written as:

$$\frac{z}{z^*} = \frac{w}{w^*}\frac{\overline{c}^*}{\overline{c}} \tag{4.3.7}$$

If the least competitive firm in an industry with a productivity draw at the upper bound of the distribution c_M has labor productivity z_M and labor is perfectly mobile between firms, equation ??? 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 ??? can be amended to yield:

$$\left(\frac{z}{z^*}\right)^k = \left(\frac{z_M}{z^*_M}\right)^k \frac{(\bar{N}/N)}{(\bar{N}^*/N^*)} \frac{1+\frac{\bar{N}*}{\bar{N}}\frac{\theta}{1-\theta}}{1+\frac{\bar{N}}{\bar{N}^*}\frac{\theta^*}{1-\theta^*}} \tag{4.3.8}$$

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 ??? combined with the expression for labor productivity gives:

$$\left(\frac{z}{z*}\right)^{k+2} = \left(\frac{w}{w^*}\right)^2 \frac{L}{L^*}\left(\frac{z_M}{z^*_M}\right)^k \frac{1-\frac{\theta}{1-\theta}}{1-\frac{\theta^*}{1-\theta^*}} \tag{4.3.9}$$

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

4.3.1 Further Changes (work in progress!)

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 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[7]. 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 labor 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 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 4.1 gives an overview of the studies used and a glance at their respective results, showing considerable variation

[7] 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.

Table 4.1: 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 (.6) Lumber (.5) Printing (.49)	Leather (.29) Food Processing (.24) Tobacco (.21)
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 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 60th 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: Food, Beverages and Tobacco, Textiles as well as Chemicals, Fuels, Plastics, Rubber fall into the long run category, Machinery and Transportation Equipment, Manufacturing not else classified and Wood and Cork make up the short run category. The three remaining industries (Paper and Printing, Basic and Manufactured Metal Products and Other Non-Metallic Mineral Products) are too close to the median to be classified either way and are thus dropped from the sample[8].

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

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

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

[8]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

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:

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

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 rho decreases $1-\rho$ and thus increases $\bar{p}$). Now assume further, that price setting decisions and substitution behavior 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 behavior. 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.

4.4 Application

Starting from the equations for prices, productivity and markups derived within the Melitz-Ottaviano framework, Chen et al. then proceed by constructing log-linearized equations to estimate. The estimation equation for prices is given by:

$$\begin{aligned}\Delta \ln\left(\frac{\bar{p}_{it}}{\bar{p}^*_{it}}\right) =& \beta_0 + \beta_1 \Delta \ln \theta_{it} + \beta_2 \Delta \ln \theta^*_{it} + \beta_3 \Delta \ln D_{it} + \beta_4 \Delta \ln D^*_{it} \\ &+ \gamma \left[\ln\left(\frac{\bar{p}_{it-1}}{\bar{p}^*_{it-1}}\right) + \delta_0 + \delta_1 \ln L_{t-1} + \delta_2 \ln L^*_{t-1} + \delta_3 \ln \theta_{i,t-1} + \delta_4 \ln \theta^*_{i,t-1}\right] + \varepsilon_{ijt}\end{aligned} \tag{4.4.10}$$

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 openness lowers relative prices in the short-run, and hence $\beta_2 > 0$ as well as $\beta_3 < 0$ and $\beta_4 > 0$ – a higher number of domestic producers lowers the domestic price level while the number of foreign producers has the opposite effect.

The model specification used in this analysis is quite similar to that of Chen et al. (2009), however, we have included an arguably better measurement of domestic openness: tariffs imposed on foreign products. The estimation equation for prices utilized in this analysis is given by:

$$\begin{aligned}\Delta \ln\left(\frac{\bar{p}_{it}}{\bar{p}^*_{it}}\right) =& \beta_0 + \beta_1 \Delta\tau_{it} + \beta_2 \Delta\tau^*_{it} + \beta_4 \Delta \ln D_{it} + \beta_5 \Delta \ln D^*_{it} \\ & + \gamma\left[\ln\left(\frac{\bar{p}_{it-1}}{\bar{p}^*_{it-1}}\right) + \delta_0 + \delta_1 \ln L_{t-1} + \delta_2 \ln L^*_{t-1} + \delta_3 \tau_{i,t-1} + \delta_4 \tau^*_{i,t-1}\right] + \varepsilon_{ijt}\end{aligned} \tag{4.4.11}$$

Similarly, the estimation equation that is utilized in this analysis for markups is given by:

$$\begin{aligned}\Delta \ln\left(\frac{\bar{\mu}_{it}}{\bar{\mu}^*_{it}}\right) =& \beta_0 + \beta_1 \Delta\tau_{it} + \beta_2 \Delta\tau^*_{it} + \beta_3 \Delta \ln \theta_{it} + \beta_4 \Delta \ln \theta^*_{it} + \beta_5 \Delta \ln D_{it} + \beta_6 \Delta \ln D^*_{it} \\ & + \gamma\left[\ln\left(\frac{\bar{\mu}_{it-1}}{\bar{\mu}^*_{it-1}}\right) + \delta_0 + \delta_1 \ln L_{t-1} + \delta_2 \ln L^*_{t-1} + \delta_3 \tau_{i,t-1} + \delta_4 \tau^*_{i,t-1}\right. \\ & \left. + \delta_5 \ln \theta_{i,t-1} + \delta_6 \ln \theta^*_{i,t-1}\right] + \varepsilon_{ijt}\end{aligned} \tag{4.4.12}$$

The effect of tariffs, openness, number of firms and market size on productivity is estimated by:

$$\begin{aligned}\Delta \ln\left(\frac{\bar{z}_{it}}{\bar{z}^*_{it}}\right) =& \beta_0 + \beta_1 \Delta\tau_{it} + \beta_2 \Delta\tau^*_{it} + \beta_3 \Delta \ln \theta_{it} + \beta_4 \Delta \ln \theta^*_{it} + \beta_5 \Delta \ln D_{it} + \beta_6 \Delta \ln D^*_{it} \\ & + \gamma\left[\ln\left(\frac{z_{it-1}}{z^*_{it-1}}\right) + \delta_0 + \delta_1 \ln L_{t-1} + \delta_2 \ln L^*_{t-1} + \delta_3 \tau_{i,t-1} + \delta_4 \tau^*_{i,t-1}\right. \\ & \left. + \delta_5 \ln \theta_{i,t-1} + \delta_6 \ln \theta^*_{i,t-1} + \delta_7 \ln w_{i,t-1} + \delta_8 \ln w^*_{i,t-1}\right] + \varepsilon_{ijt}\end{aligned} \tag{4.4.13}$$

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.

4.4.1 Dataset

The database we utilize covers the period 1994-2006 for NAFTA member countries – Canada, Mexico, and the US – and nine manufacturing sectors. 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. While the majority of the manufacturing data collected is reported at the two-digit level according to ISIC Rev. 3, we aggregate all data according to ISIC Rev. 2 in order to keep consistency throughout the analysis (see Table 4.4 for the manufacturing classification breakdown).

The value of markups is not easily observable, and thus we follow the calculation as outlined in the recent literature on industrial organization, and similarly used by Chen et al. (2009), although with a subtle modification. We compute average markups as the ratio of sectoral turnover relative to total variable costs, which are computed as the sum of intermediate inputs and labor costs as reported by the OECD STAN database.[9] Due to the unavailability of data on sectoral turnover, we use sectoral production as a proxy. While turnover may be slightly higher than production[10] in a given year if all of the produced goods are sold along with any stored goods from previous years, according to the OECD these measures will converge in the long term. Fixed costs are excluded from the calculation, as they will cause a negative bias between markups and openness. As the number of foreign firms increases due to a decrease in trade costs, the market share for domestic producers

[9] In contrast, Chen et al. (2009) calculate average markups as the ratio of turnover relative to the sum of the costs of materials, consumables, and staff costs.

[10] Production represents the value of goods produced in a year, whether they are sold or stocked.

falls; because this will spread fixed costs across a smaller share of production, average total costs will increase contemporaneously (Chen et al., 2009).

Labor productivity is calculated as the ratio between real value added and total employment. Value added is reported by the OECD STAN database, while employment data is provided by LABORSTA, the statistical department of the International Labor Organization (ILO).

The main explanatory variable in this analysis is the tariff imposed on foreign products. The simple average for tariffs disaggregated at the two-digit level (ISIC Rev. 2) imposed on the bi-lateral trading partners within NAFTA were taken from the World Integrated Trade Solution (WITS), which is a resource developed by the World Bank. The changes in import tariffs across all industries for each country-pair are illustrated in Figures 4.1 – 4.6, including the maximum and minimum tariff for each year. The tariffs imposed by the U.S. on Canadian and Mexican goods, and those imposed by Canada on Mexican and U.S. goods approached zero by 1999–2000, while the Mexican imposed tariffs fluctuated drastically throughout the time period, with some tariffs reaching 30%. The average, maximum, and minimum tariffs by country and sector are also detailed in Table 4.3, as well as the other main variables in this analysis.

To construct the openness variable, we calculate the ratio of imports relative to the sum of imports and domestic production, as in Chen et al. (2009). The bilateral trade data for each country-pair comes from the OECD STAN database and measured in current USD. Domestic production is calculated as the difference between total sectoral output and exports. The sectoral output data are also taken from the OECD STAN database (measured in current units of national currency) and converted to current USD using the OECD-reported exchange rates. In order to check the reliability of our measure of openness, we also use the UPenn openness index.

As discussed in the previous section, for all of the log-linearized 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; for this reason, the coefficient on this variable will always be at the upper bound and must be evaluated with caution. The domestic establishment data is made available by CANSIM, the U.S. Bureau of Labor Statistics, and the Instituto Nacional de Estadistica.

The control variables used in the regressions include nominal wages and size of the economy, both of which are provided by LABORSTA. Wages are disaggregated according to ISIC Rev. 3 and reported in national currency units per hour for all wage earners (as before, we convert all data into ISIC Rev. 2 and current USD). For the size of the economy L, we utilize the total economically active population for manufacturing. Moreover, there are missing values for the US and Mexico[11] and thus we use interpolated data for these years.

4.4.2 Estimation

As outlined at the beginning of Section 4.4, 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

[11] We use interpolated data for the years 1999, 2001, and 2003 for the US, and 1992, 1994, 1996, 1998, 2001-2003 for Mexico.

price, markups, and labor 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. Moreover, because our variables are stationary in a unit root sense and serially correlated, we utilize a panel-specific AR(1) autocorrelation structure and perform our regressions using a Generalized Least Squares estimation strategy.

Table 4.2 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.

4.5 Results & Discussion

Tables 4.5, 4.6 and 4.7 present our results on the short-run effects of trade liberalization on prices, markups and productivity, respectively. Column (1) in each table presents the results from our theoretical estimations in equations 4.4.11, 4.4.12,

Table 4.2: Comparative Statics – Model Predictions

Regressor	**Dependent Variables**					
	P_{sr}	P_{lr}	μ_{sr}	μ_{lr}	z_{sr}	z_{lr}
τ	$+$	$-$	$+$	$-$	$-$	$+$
τ^*	$-$	$+$	$-$	$+$	$+$	$-$
θ	$-$	$+$	$-$	$+$	$+$	$-$
θ^*	$+$	$-$	$+$	$-$	$-$	$+$
D	$-$	$+$	$-$	$+$	$+$	
D^*	$+$	$-$	$+$	$-$	$-$	
L						$+$
L^*						$-$

and 4.4.13, respectively. Beginning with table 4.5, we see that the signs on the tariff measures are as predicted – a decrease in the domestic tariff will decrease the relative price in the short-run – but only the foreign tariff measure is significant. The measure of openness, on the other hand, is as predicted in the model and both coefficients are significant at the 1% level. Finally, contrary to the model's predictions the number of firms serving the domestic market does not have any effect on prices. The results for this model specification on markups are similar, with the openness measure as predicted and significant and no effect from the change in domestic firms (see Table 4.6); in this case, however, the change in the domestic tariff is positive as predicted and significant at the 10% level, whereas the foreign tariff rate has the opposite effect than the model predicts, albeit insignificant. Finally, Table 4.7 shows the effects on industry productivity with different results than predicted by the model. Although the domestic tariff measure is positive and significant, we see that the openness measures have the opposite effect than the model predicts and are highly significant; an increase in domestic import penetration actually brings about a *decrease* in productivity in the short-run.

Columns (2) and (3) are modified to include the log first-difference in relative tariffs and openness, as well as the log first-difference in the lagged dependent variable.

In the short run, the signs for all short–run effects of tariff reductions are as predicted and highly significant. Table 4.5 presents the effects of tariff reductions on prices in the short run. A decrease in domestic tariffs causes a decrease in relative prices, whereas foreign tariffs actually increase prices. The openness variables that are the main explanatory variable in Chen et al. (2009) are not significant when the tariffs are included, as the tariff provides a more accurate measure of openness between the country-pair; it should be noted, however, that the signs on the openness variable follow the predicted signs from the model when tariffs are excluded from the regression analysis.

Similar to relative prices, Table 4.6 focuses on price markups, and suggests that a decrease in domestic tariffs will likewise decrease domestic markups, with profit margins actually increasing with foreign tariffs. While the coefficients are no longer significant, the signs still match the predicted theory and hold with the different model specifications. In contrast, Table 4.7 focuses on labor productivity, and suggests that domestic tariff reductions will increase productivity, with foreign tariffs causing the opposite effect.

The main contribution of this paper comes from the long–run effects of trade liberalization as presented in Tables 4.8, 4.9, and 4.10. Table 4.8 focuses on the long–run effects of liberalization on relative prices. While the sign on the coefficient in question (i.e., lagged domestic tariffs) is positive, which is in direct contrast to the theory of Melitz and Ottaviano (2009), the results are not significant. However, the market size (L) has a positive and highly significant effect on relative prices in the long–run, thus contradicting theory that the most productive firms will relocate to a less competitive market and the short–run gains from trade will converge back to their equilibrium levels. As firms leave the market, relative prices remain low according to this analysis.

While the sign on tariffs was in contrast to the theory, the competitive effect of trade openness on relative markups (as presented in Table 4.9) does follow the theory of Melitz and Ottaviano, with decreasing tariffs causing an increase in relative markups in the long–run. Moreover, now the size of the market (L) has a negative impact on markups, which supports the anti-competitive effects in the long–run when firms have the option to relocate to a less competitive market.

Finally, Table 4.10 presents the results for the long–run effects of tariff reductions on labor productivity. Similar to the theoretical predictions, labor productivity is reversed in the long run.

While the estimation results of Chen et al. on country-pairs within the European Union followed those predicted by the model of Melitz and Ottaviano, the long–run results in this analysis were in contradiction to the theory with regards to prices; however, the models predictions do hold for relative markups and productivity. The estimation results for the short–run follow the model's predictions and are highly significant.

4.6 Conclusion

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. However, largely due to the year restrictions in their database, the results from Chen et al. (2009) are insignificant in the long run, and thus merely suggestive. While this paper replicates their analysis with only minor conceptual changes, our use of NAFTA as a defining multilateral trade agreement and the 12-years of data available after its implementation, provides new insights into the long–run dynamics of international trade aggrements. While the theory suggests a reversal of competitive effects in the aggregate, this paper illustrates that the competitive effects are merely assuaged, but certainly not reversed. This adds credence to the pro-liberalization camp of policy makers who argue for the social welfare improving competitive effects of trade liberalization in the short–, medium–, and long term.

4.7 Appendix A: Figures, Summary Statistics, Results

Figure 4.1: Canadian Tariff on Mexican Goods

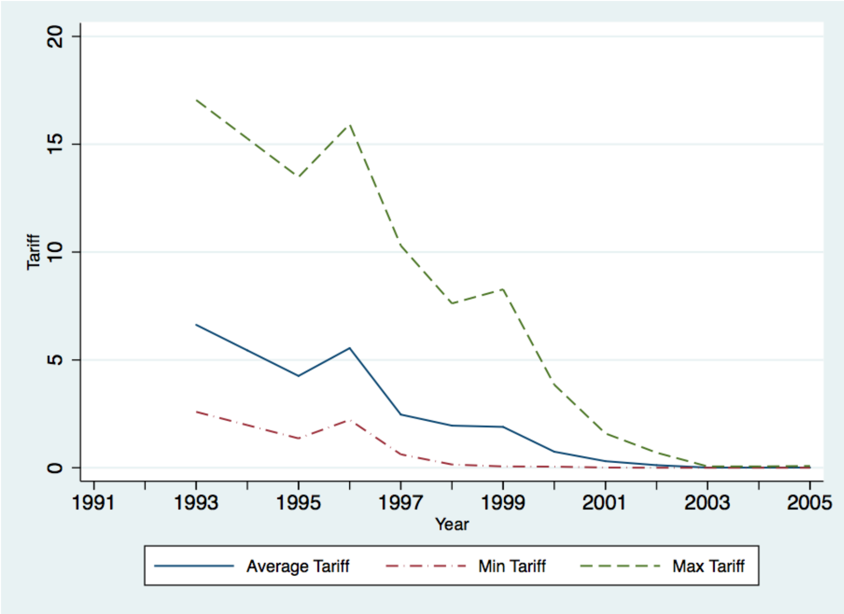
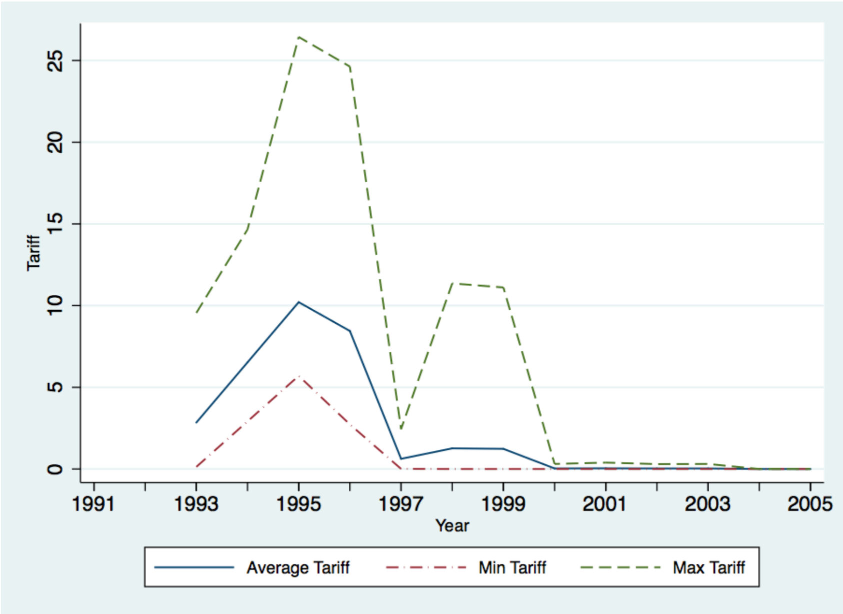


Figure 4.2: Canadian Tariff on U.S. Goods



4.8 Appendix B: Conversion Table

4.9 Appendix C: Regression Results

Figure 4.3: Mexican Tariff on Canadian Goods

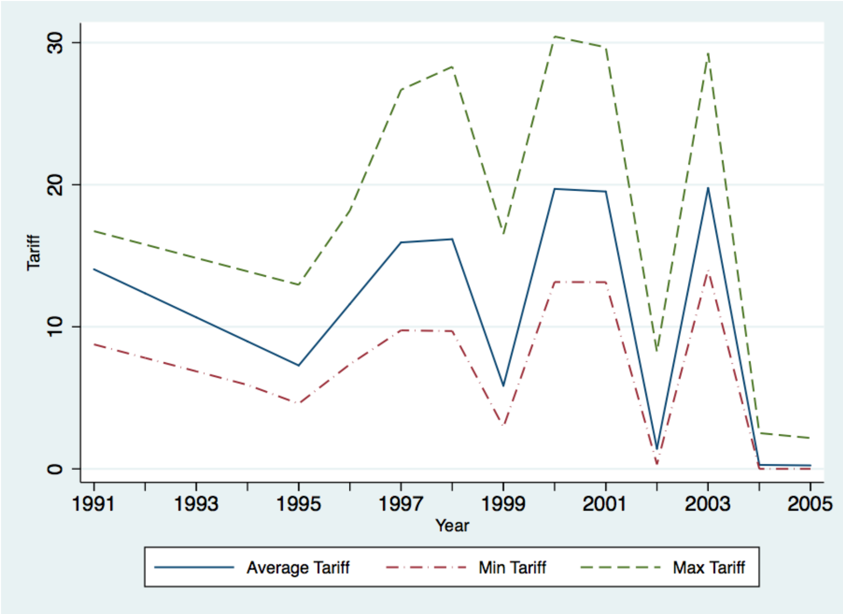


Figure 4.4: Mexican Tariff on U.S. Goods

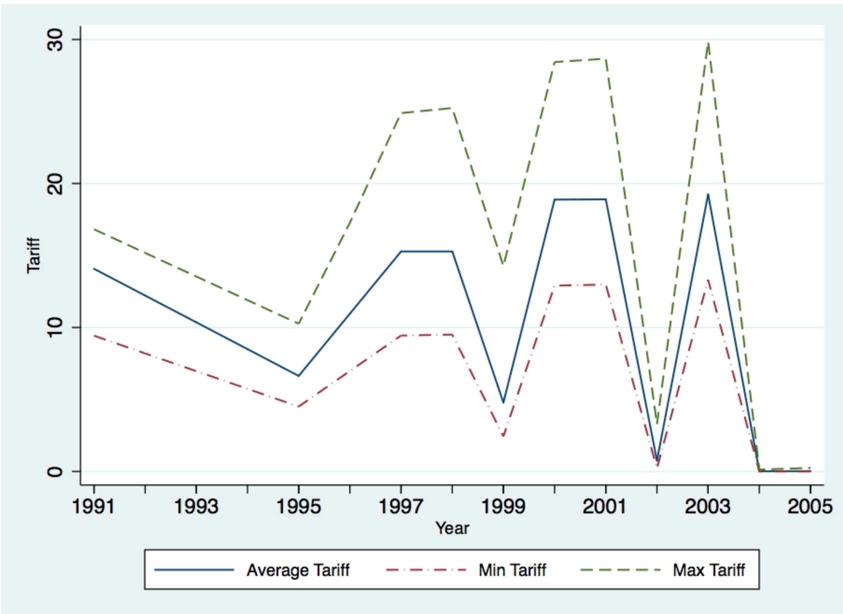


Figure 4.5: U.S. Tariff on Canadian Goods

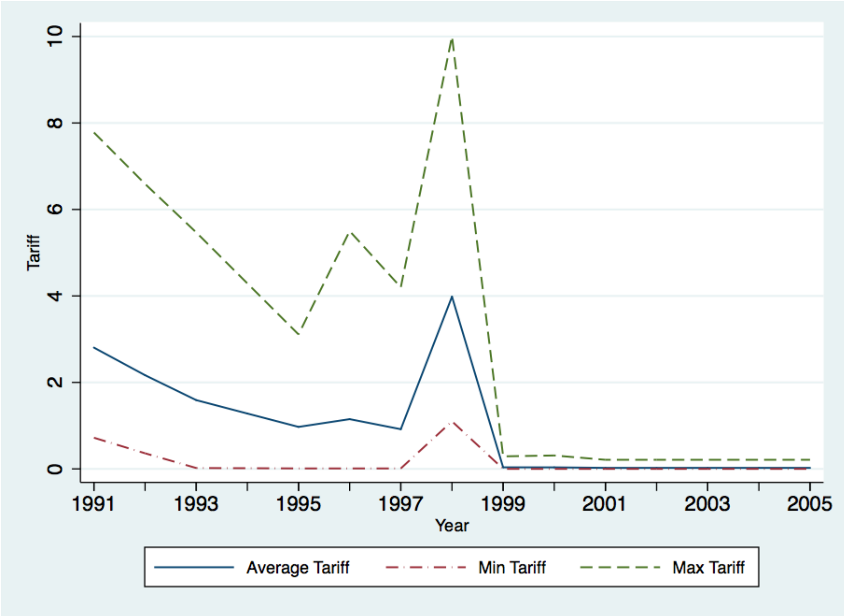


Figure 4.6: U.S. Tariff on Mexican Goods

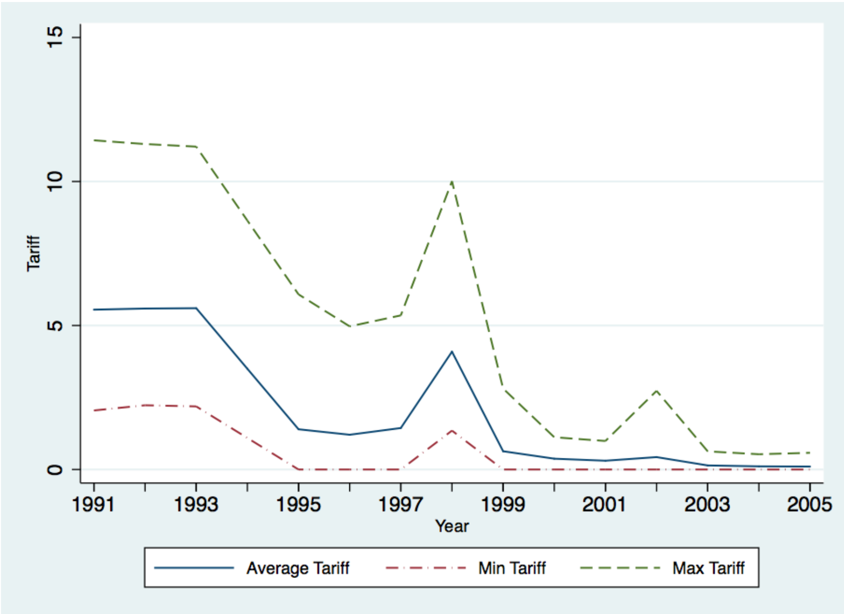


Table 4.3: Summary Statistics

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

Table 4.4: Manufacturing Sectors: ISIC Rev. 3 to ISIC Rev. 2 Conversion

ISIC Rev. 3	**ISIC Rev. 2**
D15: Foods, products, and beverages D16: Tobacco products	31: Food, beverages, and tobacco
D17: Textiles D18: Wearing apparel, dressing and dyeing of fur D19: Tanning and dressing of leather, manufacture of luggage, handbags, saddlery, harness and footwear	32: Textile, wearing apparel, and leather industries
D20: Wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	33:Wood and Wood Products, Including Furniture
D21: Paper and paper products D22: Publishing, printing, and reproduction of recorded media	34: Paper and paper products, printing and publishing
D23: Coke, refined petroleum products and nuclear fuel D24: Chemicals and chemical products D25: Rubber and plastics products	35: Chemical, petroleum, coal, rubber and plastic products
D26: Other non-metallic mineral products	36: Non-metallic mineral products, except products of petroleum and coal
D27: Basic metals	37: Basic Metal Industries
D28: Fabricated metal products, except machinery and equipment D29: Machinery and equipment n.e.c. D30: Office, accounting and computing machinery D31: Electrical machinery and apparatus n.e.c. D32: Radio, television and communication equipment and apparatus D33: Medical, precision and optical instruments, watches and clocks D34: Motor vehicles, trailers and semi-trailers D35: Other transport equipment	38: Fabricated metal products, machinery and equipment
D36: Manufacture of furniture; manufacturing n.e.c. D37: Recycling	39: Other Manufacturing Industries

Table 4.5: Prices (Short Run), all country pairs

	Dependent variable:	
	$\Delta \log \left(\frac{p}{p^*}\right)$	
	(1)	(2)
$\Delta \log \tau_t$	0.002	0.002
	(0.002)	(0.002)
$\Delta \log \tau_t^*$	-0.003^{*}	-0.003
	(0.002)	(0.002)
$\Delta \log \theta_t$	-0.106^{***}	-0.104^{***}
	(0.014)	(0.015)
$\Delta \log \theta_t^*$	0.247^{***}	0.248^{***}
	(0.023)	(0.024)
$\Delta \log D_t$	0.015	0.016
	(0.025)	(0.025)
$\Delta \log D_t^*$	-0.034	-0.084
	(0.134)	(0.153)
Observations	320	320
R^2	0.371	0.379
Note:	*p<0.1; **p<0.05; ***p<0.01	
	Fixed effects for country pair	

Table 4.6: Markups (Short Run), all country pairs

	Dependent variable:			
	$\Delta \log \left(\frac{\mu}{\mu^*}\right)$			
	(1)	(2)	(3)	(4)
$\Delta \log \tau_t$	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
$\Delta \log \tau_t^*$	0.00003	0.00003	0.0002	0.0002
	(0.001)	(0.001)	(0.001)	(0.001)
$\Delta \log \theta$			-0.024^{***}	-0.024^{***}
			(0.005)	(0.006)
$\Delta \log \theta^*$			0.028^{***}	0.028^{***}
			(0.009)	(0.009)
$\Delta \log D_t$	-0.003	-0.002	-0.005	-0.003
	(0.009)	(0.009)	(0.009)	(0.009)
$\Delta \log D_{t-1}$	0.049	0.006	0.037	-0.008
	(0.050)	(0.056)	(0.049)	(0.055)
Observations	306	306	302	302
R^2	0.009	0.006	0.109	0.107
Note:			*p<0.1; **p<0.05; ***p<0.01	
			Fixed effects for country pair	

Table 4.7: Productivity (Short Run), all country pairs

	Dependent variable:			
	$\Delta \log \left(\frac{z}{z^*}\right)$			
	(1)	(2)	(3)	(4)
$\Delta \log \tau_t$	0.007	0.007	0.009**	0.009**
	(0.005)	(0.006)	(0.004)	(0.004)
$\Delta \log \tau_t^*$	−0.006	−0.006	−0.003	−0.003
	(0.005)	(0.005)	(0.004)	(0.004)
$\Delta \log \theta_t$	−0.415	−0.389		
	(0.264)	(0.283)		
$\Delta \log \theta_t^*$	0.767*	0.681		
	(0.425)	(0.482)		
$\Delta \log D_t$			−0.329***	−0.322***
			(0.035)	(0.036)
$\Delta \log D_t^*$			0.656***	0.660***
			(0.057)	(0.059)
diff(log(firms_h))			−0.012	−0.012
			(0.061)	(0.062)
diff(log(firms_f))			0.428	0.359
			(0.333)	(0.380)
Observations	324	324	320	320
R^2	0.027	0.023	0.428	0.432
Note:			*p<0.1; **p<0.05; ***p<0.01	
			Fixed effects for country pair	

Table 4.8: Prices (Long Run), all country pairs

	Dependent variable:			
	$\Delta \log \left(\frac{p_t}{p_t^*}\right)$			
	(1)	(2)	(3)	(4)
$\Delta \log \tau_t$	-0.0004	-0.001	0.0003	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
$\Delta \log \tau_t^*$	-0.002^{*}	-0.002	-0.003^{**}	-0.003^{*}
	(0.001)	(0.001)	(0.001)	(0.001)
$\Delta \log \theta$			-0.048^{***}	-0.049^{***}
			(0.011)	(0.012)
$\Delta \log \theta^*$			0.040^{*}	0.071^{**}
			(0.020)	(0.033)
$\Delta \log D_t$	0.013	0.031^{*}	0.011	0.043^{**}
	(0.016)	(0.019)	(0.016)	(0.018)
$\Delta \log D_t^*$	-0.040	-0.031	-0.025	-0.030
	(0.091)	(0.104)	(0.091)	(0.106)
$\log \left(\frac{p_{t-1}}{p_{t-1}^*}\right)$	-0.279^{***}	-0.263^{***}	-0.270^{***}	-0.233^{***}
	(0.011)	(0.013)	(0.014)	(0.021)
$\log \tau_{t-1}$	0.001	0.001	0.002^{*}	0.001
	(0.001)	(0.001)	(0.001)	(0.002)
$\log \tau_{t-1}^*$	-0.002^{*}	-0.002	-0.003^{***}	-0.002
	(0.001)	(0.002)	(0.001)	(0.002)
$\log \theta_{t-1}$			-0.012^{**}	-0.005
			(0.006)	(0.011)
$\log \theta_{t-1}^*$			0.012^{**}	0.017
			(0.006)	(0.025)
$\log D_{t-1}$	0.0004	0.040^{**}	0.003	0.069^{***}
	(0.005)	(0.020)	(0.005)	(0.021)
$\log D_{t-1}^*$	0.001	0.088^{*}	-0.0005	0.082^{*}
	(0.004)	(0.045)	(0.004)	(0.045)
$\log L_{t-1}$	-0.082	-0.175	-0.135	-0.323^{**}
	(0.150)	(0.156)	(0.148)	(0.160)
$\log L_{t-1}^*$	0.061	0.132	0.128	0.251
	(0.152)	(0.153)	(0.151)	(0.154)
Observations	324	324	320	320
R^2	0.731	0.751	0.753	0.776
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$ Fixed effects for country pair

Table 4.9: Markups (Long Run), all country pairs

	Dependent variable:			
	$\Delta \log \left(\frac{\mu}{\mu^*}\right)$			
	(1)	(2)	(3)	(4)
$\Delta \log \tau$	0.001	0.0004	0.001	0.0005
	(0.001)	(0.001)	(0.001)	(0.001)
$\Delta \log \tau^*$	0.001	0.001	0.00002	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)
$\Delta \log \theta$			-0.032^{***}	-0.027^{***}
			(0.006)	(0.006)
$\Delta \log \theta^*$			0.048^{***}	0.095^{***}
			(0.010)	(0.016)
$\Delta \log D_t$	-0.0004	0.012	-0.004	0.006
	(0.009)	(0.008)	(0.009)	(0.007)
$\Delta \log D_t^*$	0.027	-0.055	0.034	-0.063
	(0.053)	(0.050)	(0.051)	(0.047)
$\log \left(\frac{\mu_{t-1}}{\mu^*_{t-1}}\right)$	0.010	-0.460^{***}	0.003	-0.454^{***}
	(0.007)	(0.042)	(0.007)	(0.039)
τ_{t-1}	-0.001	-0.002^{**}	-0.001	-0.002^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
τ^*_{t-1}	0.001	0.001^{*}	-0.0001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
$\log \theta_{t-1}$			-0.010^{***}	-0.010^{*}
			(0.003)	(0.005)
$\log \theta^*_{t-1}$			0.008^{**}	0.037^{***}
			(0.003)	(0.010)
L_{t-1}	-0.156^{*}	-0.128^{*}	-0.255^{***}	-0.185^{***}
	(0.080)	(0.068)	(0.080)	(0.066)
L^*_{t-1}	0.139^{*}	0.104	0.230^{***}	0.138^{**}
	(0.081)	(0.068)	(0.081)	(0.068)
Observations	306	306	302	302
R^2	0.040	0.332	0.183	0.491
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$	
			Fixed effects for country pair	

Table 4.10: 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

Chapter 5

Conclusion

Throughout this thesis, an outline of affine term structure models is provided. This particular class of term structure models has been made very popular in recent years due to its ability to capture the dynamics of yields both across their time series and cross-section and its ease in imposing the absence of arbitrage, allowing in turn the obtention of adaptable risk premia specifications. Affine term structure models have the advantage of allowing various extensions, in a wide range, to their basic primary setup, asserting their importance in the literature. However, difficulties do arise in their estimation and in the interpretation of the latent factors used. This thesis addresses both problems by utilizing a specific structure to the factor loadings, known as the Nelson-Siegel method. The estimation of this term structure model not only circumvents the global optimum issues but further provides some interpretation to the factors, given the level, slope and curvature factors of the Nelson-Siegel interpolation are not only intuitive in their nature, but also have reliable macroeconomic links.

The present thesis introduces and employs dynamic term structure models to macroeconomic and financial research questions. More precisely, this study initially pertains to financial markets by establishing a tie between interest rates and exchange rates. The study follows by concerning itself with macroeconomic objectives, by exploiting the relationship between yields and inflation.

In a first instance, this study exploits a theoretical relationship between interest rates and exchange rates, namely the uncovered interest rate parity, with the aim to extract currency risk premia through a bilateral affine term structure model with stochastic volatility. The method proposed consists of developing an affine Arbitrage-Free class of dynamic Nelson-Siegel term structure models (AFNS) with stochastic volatility to obtain the domestic and foreign discount rate variations, which in turn are used to derive a representation of exchange rate depreciations. The manipulation of no-arbitrage restrictions allows to endogenously capture currency risk premia. The estimation exercise comprises of a state-space analysis using the Kalman filter. The imposition of the Dynamic Nelson-Siegel (DNS) structure allows for a tractable and robust estimation, offering significant computational benefits, whilst no-arbitrage restrictions enforce the model with theoretically appealing properties. Empirical findings suggest that estimated currency risk premia are able to account for the forward premium puzzle.

In a second instance, inflation expectations and inflation risk premia are derived using a shadow rate class of term structure models. In response to the recent financial crisis, the Bank of England reduced short term interest rates to 0.5%. With such low short term rates, traditional term structure models are likely to be inappropriate for estimating inflation expectations and risk premia, because expectations based on such models might implicitly violate the zero lower bound condition. In this segment both the nominal and real UK term structure of interest rates are studied, using the dynamic term structure model introduced by **?**, which imposes the non-negativity of nominal short maturity rates. Estimates of the term premia, inflation risk premia and market-implied inflation expectations are provided. Findings indicate that the zero lower bound specification is necessary to reflect countercyclicality in nominal term premia projections and that medium and long term inflation expectations have been contained within narrower bounds since the early 1990s, suggesting monetary policy credibility after the introduction of inflation targeting.

For my future research projects, I wish to draw from the analysis and discussion of this thesis and elaborate further on this strand of the literature, this time emphasizing on the joint effect of monetary economics and finance on asset prices, financial markets and monetary policy. Two main perspectives emerge within my research agenda. A potential project, that inclines more towards macroeconomic concepts, consists in building an extension of the two above-mentioned models by constructing a Taylor rule type of model which would further extend to include growth. Furthermore, an alternative suggests to further exploit the interaction between macroeconomic and financial data to explore a gap in the literature. Specifically, the study includes providing an economic interpretation to the latent factors, used in the state-space representation, by venturing towards macro-finance models and high frequency data. This analysis is built on the prior belief that assets are affected by macroeconomic conditions but simultaneously suffer from microstructure phenomena.

Notwithstanding the extensions listed above, it is crucial to note that the most important message to draw from this thesis is that the literature on risk premia is still at its infancy due to the striking complexity involved in estimating an unobservable variable which nonetheless contains a very rich informational content. In turn, in future research, I wish to investigate the sensitivity of the price of risk, and consequently of risk premia, to different specifications in the mean reversion matrix of the states' dynamics. The aim is to determine a preferred specification for dynamic term structure models using a Bayesian shrinkage estimation approach.

To conclude, this thesis builds a spherical account of the versatility of affine models by implementing them to distinct monetary finance applications. Several of the pending issues in the literature are addressed and the grounds for future interesting questions are paved.

Bibliography

Aguiar, Mark A. and Mark Bils, "Has Consumption Inequality Mirrored Income Inequality?," NBER Working Papers 16807, National Bureau of Economic Research, Inc February 2011. 31

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130. 46

_ , _ , Dave Donaldson, and Andres Rodriguez-Clare, "The Elusive Pro-Competitive Effects of Trade," 2012. Working Paper. 46

Attanasio, Orazio, Erik Hurst, and Luigi Pistaferri, "The Evolution of Income, Consumption, and Leisure Inequality in the US, 1980-2010," in "Improving the Measurement of Consumer Expenditures," National Bureau of Economic Research, Inc, 2013. 31

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, "Trends in U.S. Wage Inequality: Re-Assessing the Revisionists," NBER Working Papers 11627, National Bureau of Economic Research, Inc September 2005. 30

_ , _ , and _ , "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, May 2008, *90* (2), 300–323. 39

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. 56

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. 28, 33

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. 50

Bernard, Andrew and Bradford Jensen, "Exporters, Jobs and Wages in U.S. Manufacturing," *Brookings Papers on Economic Activity: Microeconomics*, 1995, pp. 67–112. 39

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. 49

Bertrand, Marianne and Adair Morse, "Trickle-Down Consumption," Working Paper 18883, National Bureau of Economic Research March 2013. 28

Brown, Mark W., "Renewing Canada's Manufacturing Economy: A Regional Comparison, 1973-1996," Technical Report 11 2004. 56

Calderon-Madrid, Angel and Alexandru Voicu, "The NAFTA Tide : Lifting the Larger and Better Boats," 2007. Working Paper. 49

Campbell, Jeffrey R. and Zvi Hercowitz, "The Role of Collateralized Household Debt in Macroeconomic Stabilization," NBER Working Papers 11330, National Bureau of Economic Research, Inc May 2005. 27

Card, David, Jrg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *The Quarterly Journal of Economics*, 2013, *128* (3), 967–1015. 39

Carroll, Christopher D., "Solving consumption models with multiplicative habits," *Economics Letters*, July 2000, *68* (1), 67–77. 35

Chen, Natalie, Jean Imbs, and Andrew Scott, "The dynamics of trade and competition," *Journal of International Economics*, February 2009, *77* (1), 50–62. 11, 12, 47, 48, 50, 53, 54, 59, 61, 63, 64, 68

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. 53

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. 50

Cordoba, Juan-Carlos, "U.S. inequality: Debt constraints or incomplete asset markets?," *Journal of Monetary Economics*, March 2008, *55* (2), 350–364. 32

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. 27, 39

DiPrete, Thomas A, "Is This a Great Country? Upward Mobility and the Chance for Riches in Contemporary America," *Research in Social Stratification and Mobility*, 2007, *25*, 89–95. 41

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. 56

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. 56

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. 49

Glaeser, Edward L., "Psychology and the Market," *American Economic Review*, May 2004, *94* (2), 408–413. 40

Gordon, Robert J., "Misperceptions About the Magnitude and Timing of Changes in American Income Inequality," NBER Working Papers 15351, National Bureau of Economic Research, Inc September 2009. 30

Grossman, Gene M and Elhanan Helpman, "Trade, Innovation, and Growth," *American Economic Review*, May 1990, *80* (2), 86–91. 46

Guvenen, Fatih, "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, June 2007, *97* (3), 687–712. 11, 28, 35, 38

Head, Keith and John Ries, "Rationalization effects of tariff reductions," *Journal of International Economics*, April 1999, *47* (2), 295–320. 49, 55

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante, "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," *Review of Economic Dynamics*, January 2010, *13* (1), 15–51. 30

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

Jantti, Markus and Stephen P. Jenkins, "Income mobility," ISER Working Paper Series 2013-23, Institute for Social and Economic Research November 2013. 39

Kambourov, Gueorgui and Iourii Manovskii, "Occupational Specificity Of Human Capital," *International Economic Review*, 02 2009, *50* (1), 63–115. 39

Kilian, Lutz, Alessandro Rebucci, and Nikola Spatafora, "Oil shocks and external balances," *Journal of International Economics*, April 2009, *77* (2), 181–194. 53

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. 27, 39

Krueger, Dirk and Fabrizio Perri, "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies*, 2006, *73* (1), 163–193. 27, 30, 32

Ma, Lin, "Globalization and Top Income Shares," Technical Report, University of Michigan 2013. Working Paper. 30

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, *71* (6), 1695–1725. 46, 49, 50, 51

_ **and Daniel Trefler**, "Gains from Trade When Firms Matter," *Journal of Economic Perspectives*, Spring 2012, *26* (2), 91–118. 46

_ **and Gianmarco I. P. Ottaviano**, "Market Size, Trade, and Productivity," *Review of Economic Studies*, January 2008, *75* (1), 295–316. 47, 50, 51

_ **and Stephen J. Redding**, "New Trade Models, New Welfare Implications," NBER Working Papers 18919, National Bureau of Economic Research, Inc March 2013. 46

Moore, David W., "Half of Young People Expect to Strike It Rich," March 2003. 41

Narajabad, Borghan Nezami, "Information Technology and the Rise of Household Bankruptcy," *Review of Economic Dynamics*, October 2012, *15* (4), 526–550. 27

Ottaviano, Gianmarco I. P., Takatoshi Tabuchi, and Jacques-Francois Thisse, "Agglomeration and Trade Revisited," *International Economic Review*, 2002, *43* (2), 409–436. 51

Piketty, Thomas and Emmanuel Saez, "Income Inequality In The United States, 1913-1998," *The Quarterly Journal of Economics*, February 2003, *118* (1), 1–39. 30

Rajan, Raghuram, *Fault Lines: How Hidden Fractures Still Threaten the World Economy*, Princeton University Press, 2011. 27

Romalis, John, "NAFTA's and CUSFTA's Impact on International Trade," *Review of Economics and Statistics*, 2007, *89* (3), 416–435. 49

Saez, Emmanuel and Gabriel Zucman, "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," NBER Working Paper October 2014. 33

Trefler, Daniel, "The Long and Short of the Canada-U.S. Free Trade Agreement," *American Economic Review*, 2004, *94* (4), 870–895. 49

van Treeck, Till, "Did inequality cause the U.S. financial crisis?," IMK Working Paper 91-2012, IMK at the Hans Boeckler Foundation, Macroeconomic Policy Institute 2012. 28