Quantifying the Distributional Implications of Trade: Occupational Reallocation in Denmark

Job Market Paper

Sharon Traiberman*

Princeton University
straiber@princeton.edu

August 1, 2015

Abstract

There is a growing concern that many workers do not share in the gains from trade. In this paper, I argue that occupational reallocation plays a crucial role in determining the winners and losers from trade liberalization: two workers in the same industry can be affected entirely differently by trade shocks, depending on what they do. Affected workers also face a period of costly adjustment. To quantify these effects, I construct and estimate a dynamic Roy model of the Danish labor market. The large number of occupations complicates estimation. To overcome this issue I project occupations onto a low-dimensional task space. This parameter reduction coupled with conditional choice probability techniques yields a tractable nonlinear least squares problem. My results suggest the costs of occupational transition can be large: on the order of ten years income. I find that switching sectors but not occupations is only one third as costly as switching occupations within a sector. Moreover, these costs rise with age, inducing older workers to exit or choose low paying occupations in the presence of shocks. Finally, I perform counterfactual labor market interventions.

1 Introduction

Free trade creates winners and losers, both in the short and long term. These distributional consequences result from economic activity shifting between industries, firms and occupations. Although recent work has explored the role of both industries and firms, we have yet to quantify the role of one's occupation in determining the uneven effects of trade shocks. However, both theory (Grossman and Rossi-Hansberg,

*This work has benefitted from the guidance, input and suggestions of numerous people. First and foremost, I thank my advisors Stephen Redding and Jan De Loecker.I have also learned a great deal from conversations with and comments from Orley Ashenfelter, Rafael Dix Carneiro, Kirill Evdokimov, Bo Honore, Richard Katzwer, Pauline Leung, Inessa Liskovich, Ezra Oberfeld, Alex Mas, Eduardo Morales, David Nagy, Esteban Rossi-Hansberg, Ricardo Reyes-Heroles, Christian vom Lehn, Frederic Warzynski and participants in the Princeton trade lunch, labor lunch and applied microeconomics workshop as well as participants of the Eleventh Danish International Economics Workshop. I am also especially thankful to Henning Bunzel for tremendous data help. Finally, I am grateful for generous funding from Princeton University, the Princeton IES Summer Fellowship and Aarhus University.

2008) and empirical evidence (Autor et al., 2014) suggest that much reallocation is intrasectoral—e.g., substitution from routine to knowledge-intensive tasks—thus ignoring the occupational dimension can lead to underestimates of the costs of liberalization. For this reason, this paper investigates the distributional consequences and dynamic costs of trade across different occupations.

In order to measure the full, dynamic, distributional impact of trade shocks, I build and estimate a model of occupational choice: In each period, workers choose their occupation weighing their menu of wages against the costs of switching occupations and the inability to transfer skills across jobs. In the model, trade shocks reduce the demand for labor, forcing workers to engage in costly readjustment. With my model I make four contribution to the literature: first, bringing to the trade literature a new, tractable method of estimating dynamic discrete choice models with large choice sets; second, quantifying heterogeneity in the effects of globalization across occupations; third, measuring the bias in adjustment costs arising from ignoring occupations; and fourth, to provide a laboratory for assessing redistributive and adjustment-oriented labor market policies.

Previewing results, I find that the costs of trade shocks can be very large depending on one's initial occupation. For example, the median cost of switching occupations is on the same order of magnitude as ten years' income. These costs grow by nearly one additional percentage point with additional year of age, suggesting particularly harsh adjustment costs for older workers. High switching costs motivate workers to move within a narrow band of similar occupations. This fact, paired with the tight correlation between wages in similar occupations, implies that trade shocks can trap workers into bouts of extended low wages.

Moreover, I find that the costs of intrasectoral movement between occupations can be very large. In fact, intrasectoral adjustment is three times costlier on average than movement across sectors, holding one's occupation fixed. Despite the steep cost, intrasectoral movement accounts for nearly half of all occupational switching. The high costs are compensated for by substantial heterogeneity in wages and unobserved benefits across occupations within sectors. These findings present a challenge for the current trade literature's focus on industrial reallocation: the substantial activity within sectors suggests that occupations play an equally important, if not more important, role in understanding worker responses to trade shocks.

In order to estimate the model, and thus switching costs, I exploit variation in different career trajectories. The intuition of my approach is simple: workers' occupational choices and rates of switching, controlling for income, reveal information about costs and benefits of changing occupations. The actual procedure is complicated by the presence of worker heterogeneity and unobserved continuation values. The latter arise as a consequence of state-dependent switching costs, which add a dynamic consideration to the worker's problem. Ignoring these complications can severely bias estimates of switching costs. First, ignoring heterogeneity ignores the fact that workers select into occupations as well as some workers face lower costs than others. Second, continuation values add a dynamic component to the workers' decision implying that income differentials are no longer a sufficient statistic for the relative value of occupations. To overcome the first issue, I exploit the EM algorithm applied to the joint empirical distribution of workers and occupational switching to non-parametrically separate workers into a finite set of types. To overcome the second issue, I exploit the fact that conditional on paying the full cost of switching workers identical on observables and unobservables who move into the same occupation face the same continuation values. This lets me compare workers who start in the same occupation, diverge for some time, but end in the same occupation. My method is related to the idea of renewal actions proposed by Arcidiacono and Miller (2011), extending the seminal work of Hotz and Miller (1993). They demonstrate formally that focusing on workers that start and end in the same effectively controls for unobservable initial conditions and continuation values.

The major advantage of my procedure is that it reduces estimation to a series of simple non-linear regressions, avoiding the need to solve the model directly. This is important for two reasons. First, the structural parameters can be solved without ever specifying workers' expectations or having to solve their dynamic problem. The problem of solving workers' expectations has a long history and typically requires one to substantially reduce the state space to accommodate an approximating procedure. My results, however, are robust to any approximation as they do not rely on specifying beliefs. Second, not needing to solve the model is necessary as the computational burden of dealing with occupations—which far outnumber sectors—demands a different approach than indirect inference, the strategy pursued by

Dix-Carneiro (2014) and others. This is because the computational burden of integrating a large number of shocks makes iterating over model solutions infeasible. Expectations and computational burdens are not specific to the occupational choice setting, nor are the existence of renewal actions. In fact, any researcher trying to model dynamics with a large choice set faces the same problems. And yet, dynamic trade models with large choice sets have grown in popularity. For example, models of geographic location as in Caliendo et al. (2015) or of location choice as in Morales et al. (2014). While the idea of estimation with renewal actions exists in the economics literature, my paper brings this idea to bear on trade questions and show how to implement them in a trade context.

However, my strategy faces one additional obstacle: the enormous size of the parameter space. This is a problem endemic to models with a large choice. For example, the number of pairwise switching costs grows quadratically in the choice set—implying over one thousand parameters. To circumvent this issue, I build on an from the IO literature, specifically Berry et al. (1995), and use survey data to project occupations onto a lower dimensional characteristic space. In past work this has involved projecting taste of goods, such as cars, onto tastes for various attributes, such as air conditioning. In my context, I think of occupations as a bundle of elementary tasks, each with a different importance. For example, machine workers spend substantial time on routine tasks involving manual dexterity and spatial acuity and less time on tasks involving mathematics or processing. By modeling occupations as a vector of characteristics, which I interpret as tasks, I can estimate switching costs by measuring and pricing differences in characteristics. For example, there is a price of moving to a more math-intensive occupation which may be perhaps mitigated if less dexterity is required. Even models with modest choice sets require stringent parametric assumptions on moving costs. My method adds to the literature by proposing a cost structure that is simultaneously parsimonious and rich—allowing for substantially more heterogeneity than the entry and exit structure pursued in Artuc and McLaren (2012), Dix-Carneiro (2014) and others.

The final benefit of my approach is that estimation of labor supply and labor demand can be decoupled. The advantage to this approach is that measurement of switching costs and other labor market parameters are robust to any particular model of labor demand that can be aggregated to economywide occupational demand curves. This allows for flexibility in specifying how trade shocks shift labor demand curves and highlights that different channels can be explored with having to reestimate labor supply. In this paper I use a model of industry spillovers. The economy is organized into broad sectors (e.g., manufacturing and services) and within these sectors are a number of highly disaggregated industries. Every industry uses some mix of occupations as well as other industry outputs as inputs. Trade shifts the price of these inputs, inducing reallocation of occupations both within and across sectors. By using a highly disaggregated IO matrix, the elasticity of substitution between imported inputs and occupations is arbitrary. Thus, even with a relatively simple production structure at the disaggregated industry level, the industry spillover model can reproduce many features of more aggregated models such as capital-skill complementarity.

I estimate my model using a linked employee-employer dataset of the Danish labor market. This setting provides several advantages over other data sets, such as the CPS. First, the dataset is large, with 2 to 2.5 million observations per year. This is crucial for estimating transitions across a large number of occupations. Second, the data is very high quality, providing details on workers' occupations and linkages to firm and industry level imports. Third, as a small open economy, Denmark allows me to focus on the direct impact of trade shocks without modeling the entire global economy. Finally, as a developed economy with a highly flexible labor market, the lessons of Denmark can be applied broadly to other developed economies. As an example, Denmark has experimented with its own set of retraining programs for displaced workers—a policy that echoes the push for more vocational training and community colleges in the United States. However, uptake of this program has been weak. My model helps understand this outcome: retraining costs, even if partially subsidized, may be too high or too difficult for actual displaced workers, inducing labor market exit rather than adjustment.

The high costs of occupational mobility raises a serious question for policy makers: how can one ensure that workers adversely affected by a trade shock are at least as well off as before liberalization? The last part of my paper deals with this question by calculating the impact of labor market policies aimed at dealing with displaced workers. In particular, I perform a series of counterfactuals involving different labor market interventions. While there is an old question of how (and whether) gainers can compensate losers from trade, I focus my attention to policies that smooth the adjustment path of workers, abstracting from a fleshed out tax and transfer system. I do this because in the real world these are the kinds of policies pursued by governments. For example, in the US the Trade Adjustment and Assistance (TAA) program focuses on helping displaced workers with retraining and on the job search. In addition to retraining, adjustment packages do often include unemployment benefits or even implicit early retirement subsidization (through, for example, SSDI). I explore both of these ideas. First, I do a simple accounting exercise: holding import prices fixed at initial levels in order to properly map out the long-term consequences of trade on worker allocation. [Discuss some]. Second, I assess the costs and benefits of the kinds of policy interventions discussed above that aim to assist negatively affected workers: lowering mobility costs (through, for example, retraining programs) and a direct subsidy to early retirement. [Discuss some].

The rest of this paper proceeds as follows. I briefly summarize the Danish labor market in section 2. In particular, I argue that its status as a small open economy with one of the most flexible labor markets in Europe make it an excellent setting for my estimation. Section 3 describes how I map my highly granular data to aggregates wieldy enough for estimation. In sections 4 and 5 I outline my econometric model and the estimation strategy used to recover key parameters. In these sections I also discuss my approach's strengths and weaknesses, positioning it in the literature. I turn to my results and counterfactual experiments in section 6. The final section concludes.

2 Econometric Model and Framework

In this section I define the labor supply model that I take to the data. I break the presentation down into several pieces: first I introduce the worker's problem in full generality, including the timing of shocks and all notation; then, I describe the state space in the model, including the presence of unobserved comparative advantage, next, I describe the parametrization of occupational switching costs; then I discuss how non-employment enters the model; finally, I turn to the specific assumptions I make on shocks.

General Setup and the Worker's Problem

At the beginning of the period, a worker, indexed by i, observes some set of deterministic state variables ω (e.g., age or ability) and a set of moving cost shocks, $\boldsymbol{\epsilon}_t$. With this information in hand, she chooses an occupation for that period. After she has makes her choice but in the same period, she receives an *ex-post* productivity shock ς_{it} that affects her supply of supplied human capital in the current period. As I'll discuss below, the worker's state includes a rich vector of observable as well as unobservable elements that guide selection. Thus, the productivity shock should be thought of as an unanticipated shock to the worker's productivity after making their decision. As the length of time I consider is a year, the unanticipated shock is a reasonable way to capture determinants of income that the workers may not know at the beginning of the period but may accrue throughout the year. For example, ς_{it}, picks up the effects of unanticipated health shocks, shocks to the establishment employing the worker, or serendipitous shocks to the workers' human capital. The moving cost shock, on the other hand, reflects an idiosyncratic shift in the cost of occupations. It can be thought of as reflecting one's mood and desire for a certain occupation. For example, a worker can wake up and realize that they have always wanted to be a baker, even if they currently work as an economist. It can also reflect any shock that lowers the burden of switching, such as a promotion; likewise, it can reflect a negative shock that forces a worker to "choose" non-employment.

When making their decision, workers maximize income, supplying human capital $h_o(\omega, \varsigma)$ inelastically at a market skill price, w_{ot}. Thus, the worker's problem is only to choose her occupation (including the possibility of non-employment). This problem can be written in the following recursive formulation:

$$v_t(o, \omega_i, \boldsymbol{\epsilon}_{it}) = \max_{o' \in \mathcal{O}} C(o, o', \omega) + \rho \epsilon_{o'it} + \eta_{o'} + w_{o't} E_{-t} h_{o'}(\omega, \varsigma_{it}) + \beta E_t V_{t+1}(o', T(\omega, o'))$$

where C are moving costs, ϵ_{oit} is the moving cost shock, w_{ot} is the wage in occupation o' at period t, $h_{o'}$ is a human capital function specific to each occupation, V_{t+1} is a continuation value and $T(\omega, o')$ determines how the state evolves given the worker's choice. The reason I use both lowercase and upper case v is to be clear about integration of shocks. A lowercase v denotes the value function *after* shocks have been observed and V refers to the value function after shocks have been integrated out.[1] The expectation operator on human capital represents the worker's lack of information about their productivity while on the continuation value, the expectation operator represents the workers' forecasting problem for future skill prices. The parameter ρ prices the moving cost shock while $\eta_{o'}$ is a non-pecuniary common payoff for all workers in occupation o'.

While wages are a common feature of labor market models, moving costs are not and require some economic interpretation. In my model, the preferred interpretation of moving costs is capturing two forces. First, switching occupations involves a costly search process. Thus, C partly reflects a search cost common to workers of all stripes. Second, and more unique to the Roy model setting, when a worker moves there is often costly retraining involved. For example, becoming a baker is no easy task. Acquiring the baseline human capital necessary to perform at a new occupation requires time and often money; the moving cost function aims to price this process in full. The need to acquire new human capital is also why it is absolutely crucial that the moving cost function depend both on one's source and target occupations. For example, while becoming a baker is difficult, it is certainly easier for a confectioner or brewer to learn the requisite skills than an economist or lawyer. As I will describe in more detail below, this motivates my characteristic-based cost function.

In addition to costs, some interpretation of η is in order. At a macroscopic level, the non-pecuniary benefits capture all the determinants of occupational choice left unexplained by income differences. While unpacking these parameters is far outside the scope of this paper, they should be thought of as representing the ease or difficulty of certain occupations. For example, surveys have confirmed that workers in the clergy or education get more satisfaction out of their job than workers in sales departments or operating machines. These differences in the desirability of occupations can lead productive workers to choose lower paying but more fulfilling occupations. The occupational fixed effects in the model are a way to explicitly account for this possibility.

As a caveat, the richness I attempt to introduce in modeling comparative advantage and a large number of occupations is geared towards my focus on the impact of import competition on the distri-

[1] Keane et al. (2011) refer to this as the EMAX function and offer a complete, rigorous introduction to the DCDP framework.

bution of wages and economic activity. However, as discussed in Altonji et al. (2013), it is essentially impossible to write down a structural model of the labor market that describes all the features of the data at once. In aiming to estimate the parameters of a particular model, I am forced to ignore many well-documented aspects of workers' career paths (e.g., learning as in Farber and Gibbons (1996) or education as in Keane and Wolpin (1997)). Nevertheless, I believe I allow for those features of the labor market that matter most when considering the impacts of free trade.

State Variables and Human Capital

The worker's state ω can be partitioned into an *observable* (to the econometrician), deterministic component and an *unobservable* (to the econometrician) time-invariant component. The observable state consists of a worker's age, their current occupational tenure, and their skill level. In the econometric model these are all discrete and assumed to live in the following space:

$$age \in [25, 59]$$

$$tenure \in [0, 6]$$

$$education \in \{0, 1, 2\}$$

Notice that there are three skill types (mapped into educational attainment) and that human capital is capped at 6 years. Workers are assumed to enter the work force at 25 and retire at 60.

The worker's *unobservable* state is a vector of time-invariant comparative advantage shocks denoted by $\boldsymbol{\theta}$. These unobservable shocks allow for rich patterns of selection on the parts of workers. In particular, it allows for workers identical to the econometrician, to nevertheless differ in their aptitude at different occupations. For example, some workers may be more well suited to office jobs such as management or law while others are more well suited to hands on work in research laboratories or manufacturing. In a model with a large number of occupations, coarse measures of education alone could never capture these difference in workers. The model also allows for arbitrary correlation in the component of $\boldsymbol{\theta}$. This captures the intuitive idea that a worker more adept at an occupation such as research is likely to be adept at similar occupations such as teaching and perhaps less so at operating machines. Of course this is only an example, I allow the data to identify the distribution of the parameters. By focusing on a particular worker's draw of $\boldsymbol{\theta}_i$ and looking at component's across occupations θ_{io} I capture the comparative advantage a worker has in different occupations. But this parameter also allows me to capture the fact that there are absolute advantage differences even within education categories. To see this effect, one can compare the value of $\boldsymbol{\theta}$ across occupations for two different workers. If $\theta_{io} > \theta_{jo}$ then clearly worker i has an absolute advantage in occupation o; moreover if $\theta_{io} > \theta_{jo}$ for all or many o then clearly worker j has an overall higher productivity than worker j.

The exact specification for the human capital function is assumed to be log-linear, in the style of a Mincer regression. This is both parsimonious and, as I will show, allows me to divide estimation into two stages. The exact expression is given by

$$h_o(\omega_{it}, \vartheta_{it}) = \exp\left\{\beta_1^o \times age_{it} + \beta_2^o \times age_{it}^2 + \beta_3^o \times ten_{it} + \beta_4^o \times \mathbf{1}\{skill_i = med\} + \beta_5^o \times \mathbf{1}\{skill_i = high) + \theta_{oi} + \sigma_o \vartheta_{iot}\right\}$$

where, recall that, ς_{it} is the worker's ex-post productivity shock. The first 5 parameters in this equation are standard covariates: a quadratic function in age, human capital returns and differential returns to skill. Notice that I allow all of these to differ by occupation. This captures that labor market experience, proxied by age, as well as occupation-specific experience, matter more in some occupations than in others. This is key for capturing selection effects that occur because of the life cycle. In particular, it allows for certain occupations to become attractive only after a worker has gained sufficient labor market experience. The θ parameters reflect unobservable absolute and comparative advantage. As discussed above, they capture absolute advantage as the mean level can vary across types and they capture comparative advantage because they can take different values for different types of workers. Including these unobservable parameters presents a challenge for estimation but are absolutely crucial to matching workers' transition patterns and also understanding how workers respond to shocks.

All that remains to be described in terms of the state and human capital is its evolution. Age increases deterministically, as does occupational tenure conditional on continued employment. I treat θ as constant over the lifecycle. While in principle I could allow for a Markov chain on the unobservable state, it is difficult to identify this underlying process given my panel length. As I'll explain later, my method can allow for shifts in unobservables, and thus time-varying changes in workers' selection criteria, provided rich enough data exists. For workers that switch occupations, I assume that conditional on paying the full cost of changing occupations, their is no additional transferability of occupation-specific human capital. Another way to say this is that two workers, identical in age, skill and type, who start in different occupations but move to a new occupation have the same starting level of human capital once they have paid retraining costs. The catch is that the cost of attaining a baseline level of human capital varies by initial occupation. Moreover, one can allow for this cost to depend on one's current talents and other state variables. Continuing with the previous section's example, I assume that while a confectioner and an economist may face different costs of learning to bake, they are equally skilled bakers at the outset, conditional on other observables. This way of modeling one's occupational tenure is imperfect as there may be some residual effects of experience I cannot capture. An alternative, pursued for example in Dix-Carneiro (2014), is to explicitly track the worker's history for some finite length T. While I can accommodate such a modeling choice, I do not pursue this for three reasons. First, with such a large state space it places extreme assumptions on the data as few workers have exactly identical long term career trajectories. Second, my focus is on occupational switching costs and estimating these precisely but also richly. It is difficult to precisely, separately identify a very large set of parameters governing experience, selection and switching costs. And while previous work has economized on parameters by sacrificing flexibility in moving costs and focusing on small choice sets, I have opted to build a rich cost side that allows heterogeneity in source occupation, target occupation as well as one's state. Third and finally, as my cost function is task-based, I interpret occupation specific human capital as learning by doing. However, when workers change occupations, once they've learned new tasks they are equally adept at performing them.

Moving Costs

I model moving costs by pricing the distance between the occupational characteristics described above. Recall that I project occupations onto a characteristic space that maps each occupation into a vector $v_o \in$

$\mathbb{R}^L$ where L is the number of characteristics. In my specification I posit the following multiplicatively separable form for the cost function:

$$C(o, o', \omega) = f(\omega)C(o, o')$$

The first function, $f(\omega)$, is an occupation-invariant, inverse moving productivity. Essentially, this captures how quickly workers can pick up new skills and the lifetime monetary costs of moving (e.g., family, geography or perhaps unobserved savings). In the estimated model I allow f to vary by age, skill, and type. This allows for heterogeneity in switching costs across workers moving from the same source occupation to the same target occupation. For example, older workers may find switching harder than younger workers, while less skilled workers may struggle relative to more skilled workers. There is also an unobservable component to switching costs which, notationally, I include as part of the vector $\boldsymbol{\theta}$. This unobservable component allows for some workers to be more adept at career changes than others. For example, there may be "quick-learners" in the population who find changing occupations easier, all else held equal. On the other hand, there may be some unmodelled risk aversion or aversion to change that makes switching harder for a subset of the population. In terms of the actual functional form I use a log-linear specification as it seems natural, but there is nothing that dictates its use:

$$f(\omega_{it}) = \exp\left\{\alpha_1 \times age_{it} + \alpha_2 \times age_{it}^2 + \alpha_3 \times \mathbf{1}\{skill_i = med\} + \alpha_4 \times \mathbf{1}\{skill_i = high) + \theta_{fi}\right\}$$

where $\theta_{fi} \in \boldsymbol{\theta}$. As a final point, I include a quadratic term in age as it improved the fit when I actually estimated the model.

The second function, $C(o, o')$ is a state-independent cost function across occupations. Recall that I model occupations as a vector of tasks. Thus, suppose there are $|\mathcal{V}|$ elementary tasks and each occupation is associated with a vector, $v^o \in \mathbb{R}_+^{|\mathcal{V}|}$. Tasks should be thought of as "a unit of work activity that produces output" (Acemoglu and Autor, 2011). Examples could be writing reports, communicating with colleagues or operating a CNC lathe. The actual value of v is an importance weight of a task to a particular occupation. For example, economists spend little time operating CNC lathes but substantial time communicating with colleagues. When one thinks of modeling occupations this way, the costs of switching occupations can be thought of as the cost of learning or dis-learning particular tasks. Continuing with the example of bakers, economists must learn to operate an oven or follow recipes; a confectioner on the other hand may need to learn very little. I will describe in the Data section how I observe occupational characteristics and their loadings. However, while the distance between characteristics is observed it is not the case that all characteristics are equally costly to change. Estimating these costs are at the heart of my paper and lead to the following specification for costs:

$$C(o,o') = \exp\left(\Gamma^M + \sum_{i=1}^{|v|} \Gamma_i \left(v_i^{o'} - v_i^{o}\right)\right)$$

The first term, Γ^M, captures a fixed cost of switching common across all pairs of occupations. This constant term reflects the baseline switching costs and I interpret it as picking up standard search costs as well as a baseline cost of retraining (this cannot be separately identified). The remaining Γ terms are the coefficients on linear distance in each component of the characteristics vector.[2] That Γ varies with v reflects that some tasks are easier to learn than others.

The problem of estimating pair-wise switching costs arises even for models with modestly sized choice sets. This is because the number of switching costs grows quadratically in the number of occupations. Typically the literature has economized on parameters by specifying an entry-exit structure of the form:

$$C(o,o') = \exp\left(C_o^{EXIT} + C_{o'}^{ENTER}\right)$$

This approach has several drawbacks which my method addresses. First of all, while the number of parameters grows linearly in the number occupations instead of quadratically, it can still become unwieldy in large models. Second and more importantly, the entry-exit structure allows for no pair-wise differences in costs. As an example, the relative cost of an economist becoming an accountant versus a baker and the relative cost to a confectioner of becoming an accountant versus a baker are

[2] In practice I allow the coefficients to differ depending on whether the movement in characteristics is positive or negative, reflecting that costs may differ if a worker needs to gain skills or shed them. Thus there are actually two Γ terms per characteristic. I also experimented with quadratic terms, but found they added little explanatory power.

the same in the entry-exit model. Such an approximation may be reasonable when the choice set is a small number of sectors, but becomes senseless when choice sets are highly disaggregated. Third, the characteristics approach offers economists substantial latitude in modeling costs. While I have chosen the linear form above as it fits the data well, I could easily handle higher order terms or interaction effects. Since the number of characteristics is fixed, it is much easier to keep the number of parameters to a reasonable number even as the choice set increases. This could be important for other trade models which feature large choice sets with observable characteristics, such as geography models.

Finally, I include dummies for switching sectors but not occupations and switching occupations but not sectors as additional components of C. These dummies allow me to compare my moving costs more cleanly with existing estimates in the literature. Clearly if sectors are irrelevant and my characteristics capture all salient features of an occupation these coefficients should be close to 0. On the other hand, if occupational characteristics play little role in determining movement but sectoral adjustment is costly than we expect the coefficients on tasks to be small while the coefficient on the sector dummy ought to be large.

Non-Employment

To model non-employment I construct a virtual payout as a simple quadratic function of age, skill and type:

$$w^N(\omega, o) = \beta_{skill} + \beta_\theta + \beta_a \times a + \beta_{a2} \times a^2$$

In the model I am silent on the reason for entering non-employment. That is to say, a shock that pushes an agent into non-employment could reflect a choice on the part of the worker, but also could reflect an unanticipated separation shock, maternity leave or any other reason for removing oneself from employment. Similarly, this value of non-employment reflects a variety of different forces affecting workers. Aside from capturing the very generous Danish social safety net, non-employment can pick up tastes for leisure, having a family or the value of home production relative to wages. While reducing the myriad reasons for non-employment to this simple structure may obscure some key features of the labor market, these are not my focus and so I set them aside.

To re-enter the workforce, I assume that workers pay the moving cost associated with their *most recent* occupation and an additional cost f_U. The state transitions in non-employment are the same as in employment with the following exception: I assume that workers keep their accumulated occupational tenure for one period of non-employment and then lose it. This is an *ad hoc* assumption chosen to fit the data, as most unemployment spells are either a single period or forever (i.e., early retirement). I use the notation of a superscript N to denote the value function or the transition matrix of a non-employed worker with state ω who was previously at occupation o. A superscript E, when unclear from context, denotes an employed worker with state ω at occupation o.

New Entrants and Retirees

So far I have only discussed existing workers. However, in every period an exogenous new cohort enters the labor force with distribution $N_t(\omega)$. There is also an initial cohort, distributed $N_0(\omega, o)$. While non-employment is a choice variable in this model, retirement is exogenously determined by a worker's position in the lifecycle.

The entry decision is similar to the switching decision and modeled as follows:

$$V_t^{Enter}(\omega_{it}, \boldsymbol{\epsilon}_{it}, \vartheta_{it}) = \max_o -f(\omega_{it})C_o^{Enter} + \rho\epsilon_{oit} + w_{ot}h_o(\omega_{it}, \vartheta_{it}) + \beta EV_{t+1}(o, T(\omega, o))$$

where C_o^{Enter} are parameters to be estimated.

Shocks

I mostly follow the extant literature and specify unobservables according to the following distributions:

$$\boldsymbol{\epsilon} \sim GEV(1)$$

$$\vartheta \sim \mathcal{N}(0, 1)$$

$$\boldsymbol{\theta} \sim (q_1, q_2, ..., q_K)$$

The extreme value assumption on switching cost shocks is standard and yields a closed form for the value function conditional on θ. The actual functional form implies that for each target occupation compared to the source occupation, the difference in switching costs is the familiar logistic distribution. Modeling wage shocks as Gaussian is also standard. In my data set, log wages are actually slightly fat tailed, even after trimming extreme outliers. Nevertheless, the Gaussian assumption provides a reasonable fit to the data as I am not interested in modeling the precise distribution of wages. Moreover, the properties of the Gaussian likelihood make it particularly attractive for estimation. Finally, I model the worker's unobserved type as coming from a discrete distribution with K types. In the actual estimation I estimate a separate distribution of types for each skill level — thus there are $S \times K$ possible groups of workers, where S is the number of skill types. While this assumption is certainly more restrictive than assuming that one's unobserved type is drawn from a continuous distribution, it allows for a rich and tractable model of comparative and absolute advantage. As I'll discuss in the estimation section, keeping unobserved heterogeneity finite is absolutely crucial to identification. With this last caveat in mind, I turn to a brief discussion of my data and then the model's estimation.

3 Data Description

The Danish data contains several databases that can be woven together to provide information on workers, such as income, occupation and place of employment. This breadth of coverage, covering the universe of Danish workers from 1997 to 2007, is what allows me to estimate the model. In this section I discuss those ingredients essential to estimation of the worker's dynamic decision problem. First, I briefly mention the datasets that I use and provide some summary statistics. Then I outline the two key aggregations from raw data to model inputs: (1) a mapping from highly disaggregated occupational and industry codes to a tractable number of occupations; (2) the construction of task space, which play the role of occupational characteristics. In the interest of brevity, I omit some details; however, a more complete description of the data and the methodology can be found in the data appendix.

3.1 Aggregating Occupational Codes

Denmark uses the ISCO system developed by the ILO in order to classify workers into occupations. The system's primary tenet is that "the basis of any classification of occupations should be the trade, profession or type of work performed by an individual, irrespective of the branch of economic activity to which he or she is attached or of his or her status in employment." That is to say, the system strives to classify occupations based on tasks and work activity.

At the most disaggregated 4 digit level there are nearly 1000 codes. However, in the structural estimation, I must aggregate both for computational reasons and because many occupations only employ a few workers. While my strategy can handle large choice sets, it is still limited by the need to have reliable estimates of transition probabilities. My aggregation strategy proceeds in two cuts. First I move from the ISCO 4 digit level to the ISCO 2 digit level. This leads to 24 occupations. These occupations are still quite specific and should be thought of as separating, for example, machinists, plant operators, drivers and craftsmen but not differentiating between type of machine. I actually disaggregate these codes by also crossing these codes 4 sectors: manufacturing, health and education, FIRE and other services. While my focus is primarily on occupations, I do this disaggregation for two important reasons. First, more disaggregated occupation codes are often found in only one sector. For example, workers coded at the two digit level as drivers in manufacturing are almost always fork lift operators while in services they are almost always taxi drivers. Thus, sectoral divisions are often a stand-in for more disaggregated occupational divisions. In this sense, they provide a natural bridge between highly disaggregated codes and less disaggregated codes. Second, including sectors allows me to benchmark my results against the literature. Most of the literature has focused on movement across broad sectors. By including this dimension directly in my estimation I can separately identify those costs associated with moving across sectors but keeping the same occupation, the costs of switching occupations within sectors and the combined costs. This allows for a direct test of the relevance of the occupational margin in thinking about worker adjustment: if occupations are irrelevant than only intersectoral movement should be costly. This procedure yields 38 total occupations, as many occupations only appear in one sector (for example, machinists are only present in manufacturing).

3.2 Occupational Characteristics

I model occupations as a bundles of tasks. As mentioned above, I think of tasks as abstract objects that represent a single unit of work. I assume there are a finite number of elementary tasks, $|\mathcal{V}|$, and that an occupation is a vector in $\mathbb{R}^{|\mathcal{V}|}$ that gives loadings on these tasks. As a concrete example, suppose there were three elementary tasks in the world—dexterity, communication and problem solving; then an occupation such as restaurant worker would have a high loading on communication with low weights elsewhere while an economist may have relatively high weight on the latter two tasks but not the first.

Tasks offer a way to put a metric on the space of occupations. If an occupation is a vector, $v_o \in \mathbb{R}^{|\mathcal{V}|}$, then one can measure the distance between occupations $d(v_o, v_o')$. To construct occupational characteristics, I need a notion of tasks that is observable. Following the labor literature, I use the O*NET database. The Department of Labor asks detailed questions of workers on the task content of their occupations. Workers are asked to rank the importance of a task on a scale of 1 to 5. Examples of tasks include "Active Learning," "Writing," "Equipment Maintenance," and "Assisting and Caring for Others." For each occupation the value recorded is an average across many workers. I standardize these values to quantiles and then treat them as cardinal. While there is no way to truly measure differences in subjective responses, the literature has treated these measures continuously and I do as well. Drawing together various surveys yields 128 questions covering 983 occupations. I relegate details of the questions and aggregation across occupations to the data appendix.

A large number of survey questions is almost as problematic as a large number of occupations. Reducing the dimensionality of the parameter space from the thousands to the hundreds is a huge step forward but remains unwieldy. Thus, the final step in my construction of tasks involves using state space reduction techniques to collapse the set of tasks to a small number of characteristics. To do so, I perform PCA on the matrix of survey responses across occupations. To determine the correct number of tasks, I follow the methodology and estimation strategies set forth in Bai and Ng (2002) and Stock and Watson (2002). These papers provide a discussion of how to modify standard information criteria, such as the AIC, for a PCA setting where one needs to determine the number of factors. In the appendix, I review this strategy as it applies to my problem.

Estimation yields 10 characteristics. Because the definition of characteristics is only unique up to orthogonal transformation, there is no "real" meaning to the characteristics. Nevertheless, it is often the case that in practice the estimated loadings still carry some intuitive content. Table 7.1 lists the survey questions with the highest loadings for each task.

4 Estimation

This section outlines my estimation procedure. To estimate the model I exploit the model's structure to transform estimation to a series of non-linear regressions. The full estimation procedure occurs in two stages: in a first stage, I estimate the distribution of time-invariant unobservables as well as the wage parameters; then in a second stage, I use a series of non-linear least squares regressions to extract the structural parameters. This is similar to the method of Scott (2014) building on Arcidiacono and Miller (2011). I present the estimation stages in reverse order: first, I outline the second stage procedure as if the worker's state were observed; second, I explain how I use the empirical distribution function of occupational switching and the Gaussian structure on income to estimate the income parameters as well as the distribution of unobserved heterogeneity.

When estimating these models, many authors in the labor literature first solve the model and then use either maximum likelihood or simulated method of moments to solve the model. Two major roadblocks prevent me from employing these methods. First, the large state and parameter space means that solving and simulating the model for every guess of parameters is prohibitively expensive—including unobserved heterogeneity in the Mincer regressions, my model has over 900 parameters. My two stage approach separates the parameters of the income equations, nearly 850 parameters, from the remaining structural parameters, of which there less than 80. The first stage procedure, detailed later, Second, workers in this model have to solve a very complicated forecasting problem. In particular, the decision of workers at time the beginning of the period depends on a forecast of the joint distribution of incomes across occupations and over time. Estimation procedures that require model solutions require also estimating this complicated object, or more realistically an approximation to this object which can imply several hundred more parameters. By exploiting the idea of renewal actions, the workers' forecast error will ultimately appear as the residual in a system of non-linear regressions. Once part of the regression error, the worker's expectations no longer play a role in estimation.

4.1 Estimating Structural Parameters

As I have stated above, the key to identification in my model are renewal actions: decisions that return workers to the same state. By focusing on workers who begin and end in the same state with a mediating period of divergent trajectories, I can exploit differences in the probability of these trajectories to pin down parameters. The challenge for the econometrician is finding what actions workers can take that leave them in the same state. I exploit the fact that, conditional on paying the switching cost, workers of the same type moving into the same occupation from different occupations are identical up to the unobserved, GEV(1) distributed moving cost shocks. Two assumptions about worker behavior underly this fact:

1. The worker's occupational choice is orthogonal to the *ex-post* income shock

2. Occupational switching is a *renewal* action—the new state only inherits the deterministic part of the prior state

The first assumption simply reiterates that the income shock is only realized after the worker makes a decision. I allow for rich selection on a host of observable variables and explicitly model unobservable selection. Thus, this assumption is similar to but actually much weaker than the conditional exogenous mobility assumption made in other worker selection models, such as Abowd et al. (1999) and Card et al. (2013). As an example, I allow for workers to have comparative advantage in certain occupations and to select their occupation based on this fact. I will discuss how this comparative advantage can be allowed to shift probabilistically, even though I cannot pursue this approach due to data constraint. Moreover, workers face idiosyncratic observable moving cost shocks that also dictate their choice of occupation. In fact, integrating out these selection shocks are how I related observed transition rates to the model's parameters.

The second assumption reasserts that when workers switch occupations, their occupation specific capital does not transfer. As discussed in the modeling section, this does not imply that workers cannot exploit their talents when transferring occupations. Actually the opposite is true: the model allows for switching costs to depend on one's initial occupation and their current levels of human capital. What the assumption does imply is that when comparing two workers moving into the same occupation, any differences in their occupational experience only manifests itself in the cost function and not through any persistent difference in wages. While ignoring this possible source of persistent differences is not innocuous, it is a relatively mild assumption compared to completely disallowing switching costs to vary pair-wise, as the extant literature does. As a final point and as will be clear, the method below *in principle* does not disallow one to track a worker's history for a finite length and build this explicitly into the wage function and state space. However, in practice a short panel (as in my setting) makes it difficult to observe a large number of workers with similar career trajectories for long stretches of time.

To describe the procedure in more detail, I take a step back and briefly review some terminology and basic formulas from the DCDP literature. First, I collect moving costs and incomes into a stage payoff denoted by $u_t(o, o', , \omega)$. Next I define the *inclusive value* and as:

$$D_t(\omega, o) = \sum_{o' \in \mathcal{O}} \exp\left(u_t(o, o', \omega) + \beta E_t V_{t+1}\left(T(\omega, o, o'), o'\right)\right)$$

This term plays the role of the denominator in the worker's transition probability and, as shown in Rust (1987), plays a prominent role in the analytic solution to the worker's dynamic problem. I leave the details to the technical appendix but building on this insight, Hotz and Miller (1993) show that the model-implied probability of observing a particular career path, conditional on initial state, can be written as the discounted sum of stage payoffs, the discounted sum of worker's expectation errors, the inclusive value and some unobserved future continuation value. The particular equation is given by,

$$\sum_{s=t}^{\tau} \beta^{(s-t)} \log \pi_s^{E/N(s)}(\omega_s, o_{s-1}, o_s) = \sum_{s=t}^{\tau} \beta^{(s-t)} u_s\left(\omega_s, o_s, o_{s-1}\right) + \sum_{s=t}^{\tau} \beta^{(s-t)} \zeta_s + \underbrace{E_\tau V_{\tau+1}(\omega_{\tau+1}, o_\tau) - \log D_t(\omega_s, o_{s-1})}_{\text{Problematic Terms}}$$

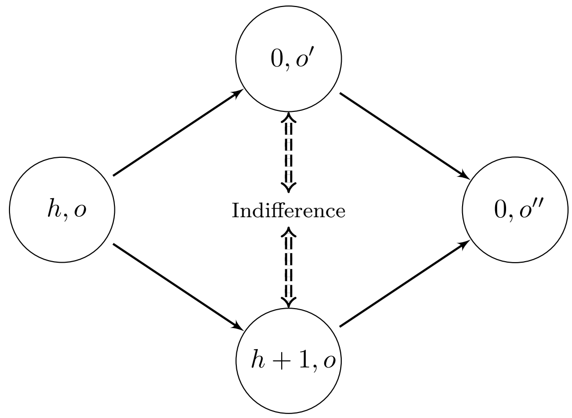
where ζ_s is the worker's forecast error on future continuation values. While the particular functional form of the equation is a result of the GEV(1) structure on shocks, the equation nevertheless can be explained intuitively. The first term, a discounted sum of stage payoffs, implies that workers are more likely to move from occupation o to o' if the actual payoff from doing so is high. The discounting implies that workers would rather wait until the future if the benefits only accrue later. The second term, a discounted sum of forecast errors, are a measure of worker optimism in the initial period. If worker's are optimistic then an econometrician will observe workers moving into an occupation at a high rate. Crucially, I assume that worker's are rational so that the expectation errors are mean zero. The last two terms I deem problematic because they are unobserved, however their interpretation is still intuitive. The first term is the worker's continuative value in τ if they make the decisions along the specified career path. The last term is the inclusive value in the initial period. Recall that this inclusive value plays a role in the solution to the worker's dynamic problem. Thus, here it plays the role of lost option value in committing to a particular career path.

To see how this equation can be used to identify structural parameters, consider taking the difference in the discounted probability of observing two different career trajectories for workers, i and j, who begin and end in the same state:

$$\sum_{s=t}^{\tau} \beta^{(s-t)} \log \frac{\pi_s(\omega_{is}, o_{is-1}, o_{is})}{\pi_s(\omega_{js}, o_{js-1}, o_{js})} = \sum_{s=t}^{\tau} \beta^{(s-t)} \left[u_s\left(\omega_{is}, o_{is}, o_{is-1}\right) - u_s\left(\omega_{js}, o_{js}, o_{js-1}\right)\right] + \tilde{\zeta}_s$$

where $\tilde{\zeta}$ is a linear combination of forecast errors. Notice that the "problematic terms" have vanished. This is the central insight of Arcidiacono and Miller (2011) and Scott (2014) that I bring into my paper. The left hand side of the above equation can be non-parametrically estimated and the right hand side is a non-linear function of observables with an additive orthogonal error term. Thus, the above equation can be used as the basis for an estimating regression.

To operationalize the above, it remains to be shown under what conditions two agents can arrive at the same state despite taking different paths (otherwise all terms would cancel to 0). This is where the notion of occupational switching as a renewal action comes into play. While there are actually many paths that would allow for the situation above to occur I consider one-shot deviations in choices. In particular, consider the sequence of choices by two individuals illustrated below:



In the example, two workers in occupation o with the same level of human capital diverge at $t+1$—one continues working in o while another goes to o'. At $t+2$ they both move to yet another occupation o''. This resets their human capital levels so that the workers are, once again, identical up to shocks. Crucially, since the initial and final value functions are the same, the workers must be indifferent to these two career paths *up to unobservable shocks.* That is to say, the only reason that the econometrician can observe the two paths above is if identical workers received different idiosyncratic, independent shocks. Thus, I can exploit deviations between observed probabilities of these paths and those implied by logit-shocks to estimate the structural parameters.

While the focus on one-shot deviations is not motivated by econometric reasons, I believe it offers a very intuitive approach to identification in the model. A drawback of model-solution based estimating strategies, such as indirect inference, is that identification can be opaque. By this I mean that pinning down the actual variation in the data that leads to variation in parameter estimates can be very difficult. However, looking at one-shot deviations in a regression framework makes the source of identifying variation clear: differences in relative transition rates, holding the differences in wages constant, is attributed to relative differences in switching costs.

To summarize the actual identifying regressions, I divide the set one-shot deviations into the following categories:

1. Early Retirement:

$$\log \frac{\pi_t^E(\omega, o, R)}{\pi_t^E(\omega, o, o)} = \alpha_s + \alpha_a \times \bar{a} - \eta_o - w_o h_o(\bar{a}, s, t) + \tilde{\zeta}_{1ot} + m_{toR}$$

2. Switching Through Employment:

$$\log \frac{\pi_t^E(\omega,o,o')}{\pi_t^E(\omega,o,o)} + \beta \log \frac{\pi_{t+1}^E(\omega,o',o'')}{\pi_{t+1}^E(\omega,o,o'')} = f(\omega)\left(C(o,o') + \beta\left[C(o',o'') - C(o,o'')\right]\right) + \left(w_{o'}h_{o'} + \eta_{o'} - w_o h_o - \eta_o\right) + \tilde{\zeta}_{2ot} + m_{too'o''}$$

3. Switching Through Unemployment:

$$\log \frac{\pi_t^E(\omega,o,R)}{\pi_t^E(\omega,o,o)} + \beta \log \frac{\pi_{t+1}^R(\omega,o,o')}{\pi_{t+1}^E(\omega,o,o')} = \beta f(\omega) f_U + (\alpha_s + \alpha_a \times a - w_o h_o - \eta_o) + \tilde{\zeta}_{3ot} + m_{toRo'}$$

where the ζ terms collect expectation errors and the m terms represent measurement error as a result of an estimated left hand side. Notice that errors are correlated in my model. While there is no serial correlation as a result of rational expectations, there is within-year correlation across sets of occupations. For example, consider two occupational switching regressions holding o, o'' and ω fixed by letting o'' differ. Then one can show that the errors will have the same expectational error terms and differ only in the measurement error (which I assume is uncorrelated across observations). To deal with this correlation I use a weighting matrix W when performing the actual regression. Because the correlation structure is very complicated and $N \approx 1.5$ million, I cannot compute an optimal weighting matrix. Instead, I use a sparse but conservative weighting matrix and relegate the details to the technical appendix.

4.2 Estimating Mincer Regressions and Unobserved Heterogeneity

In order to model unobserved heterogeneity I suppose that workers' comparative advantage is a vector θ drawn from a finite distribution $\mathbf{Q}_\Theta$. This approach is common in the structural literature and was first suggested by Heckman and Singer (1984). In the particular case of dynamic discrete choice, Crawford and Shum (2005), Dix-Carneiro (2014) and others make this assumption.[3] The particular estimation strategy I use is the Expectation-Maximization and CCP hybrid approach described in Arcidiacono and

[3] Heckman and Singer actually prove that under certain conditions, a finite distribution is the solution to a non-parametric likelihood problem for an arbitrary distribution. However, since many estimation procedures for large structural models, including mine, are not strictly maximum likelihood this proof may not hold. Instead, it is an attractive assumption from the view of computational feasibility and numerical stability.

Miller (2011). As I use their approach essentially without modification, I omit many details here and relegate the direct mapping between my model and theirs to the technical appendix. Nevertheless, I present a broad overview of the algorithm I employ.

The method begins with the likelihood function over the data including unobserved types:

$$L = \prod_{i=1}^{N} \left(\sum_{k=1}^{K} q_k \prod_{t=1}^{T} f\left(w_{it} | \omega_{it}, o_{it}, k; \Xi\right) \pi \left(\omega_{it}, o_{it} | \mathcal{H}_{it-1}, k; \Xi\right) \right)$$

where q_k is the probability of being type k, f is the Gaussian pdf on wages and π is the probability of being in state ω_{it}, o_{it} conditional on the initial state summarized by $\mathcal{H}_{t-1}$. If one could easily solve the model, then they could maximize this likelihood function to solve for unobserved states. To see how, notice that the model provides a formula for both f and π, thus allowing one to calculate the likelihood of the data. The unobserved heterogeneity present an issue by breaking the log separability of the likelihood function. However, this could be overcome with the EM algorithm.[4] Thus, if one could solve the model explicitly, he or she could use standard likelihood techniques to back out all structural parameters including the distribution of unobserved heterogeneity.

Arcidiacono and Miller's insight is that if the model can be factored into separate pieces where one piece contains some subset of model parameters and the other piece contains transition rates, then one could use the empirical likelihood function for transitions instead of the true likelihood. While this may sound complicated, it has a simple meaning in my context. Income shocks are Gaussian and independent of moving cost shocks. Thus, conditional on the observed choice of workers and their unobserved state, the likelihood for income is just the Gaussian pdf. On the other hand, the transition rates are very complicated objects. So, consider a modified likelihood given by:

$$\tilde{L} = \prod_{i=1}^{N} \left(\sum_{k=1}^{K} q_k \prod_{t=1}^{T} f\left(w_{it} | \omega_{it}, o_{it}, k; \Xi\right) \hat{\pi} \left(\omega_{it}, o_{it} | \mathcal{H}_{it-1}, k; \Xi\right) \right)$$

where the only difference between the true likelihood is the hat on the transition rates, implying that I

[4] As a brief reminder, the EM algorithm works by alternating on a guess of types (yielding an expectation) and maximizing the likelihood conditional on this guess. One updates the guess of type probabilities with each parameter update. This procedure monotonically increases the likelihood, and will converge to the maximum.

estimate non-parametrically rather than using the model-implied rates. The actual algorithm, details of which are in the appendix, now proceeds as in the standard EM algorithm. The most important piece of the algorithm is that if one specifies $\hat{\pi}$ as a linear probability models and assumes Gaussian income shocks then the entire procedure reduces to iterating on a large set of OLS regressions, allowing for a very large number of parameters to be handled tractably.

Some intuition for this method can come from thinking of this as a clustering procedure, similar to K-means clustering or a Gaussian mixture model. In both of these situations, one approximates a data generating process for a vector of observations in hopes of cutting the data into natural groups. The Gaussian Mixture Model uses the EM algorithm to exploit information embedded in a likelihood function, while K-means clustering allows one to avoid specifying a likelihood. The approach pursued above bears similarity to the mixture model approach, yielding consistent estimates of q_k since non-parametric estimates of π are consistent but does not use the full information in the likelihood. In this sense it lays between the computationally infeasible but optimal likelihood approach and a fully non-parametric clustering approach. In the econometrics literature this has a direct analog in the non-parametric dynamic discrete choice literature. In particular, Kasahara and Shimotsu (2009) discuss identification results for estimation procedures such as that use non-parametric methods in a first stage and then impose model structure in the second.

4.3 Discussion

As mentioned previously, the strength of my approach is the ability to estimate a rich model without having to explicitly model the workers' expectations. In particular, I demonstrate how even with a large state space, I can still estimate all model parameters with a series of regressions. Interestingly, my regression equations bear resemblance to the estimating equation of Artuc et al. (2010). This is because both rely on observed probabilities of movement to cancel out unobservable continuation values. However, my method extends the initial insight of Artuc et al. (2010) in four key ways. First, in recognizing the explicit role of finite dependence in the model, I am able to capture the effects of human capital and other state variables, which is a first order concern when considering the worker's dynamic problem. Their models only feature a static state space—losing the importance of the life cycle, tenure and comparative advantage. Second, the particular counterfactual histories on which I focus allow one to estimate this model without resorting to any instrumental variables. Instead, my second stage features orthogonal errors by construction. Third, in projecting moving costs onto observable characteristics I can estimate a flexible moving cost function that remains parsimonious. Finally, in tying together the likelihood and the conditional choice probabilities as in Arcidiacono and Miller, I can control for unobservable heterogeneity which enriches the model and also helps with selection problems.

Before moving to the results, I wish to discuss the bias resulting from one weakness of the model: the inability for workers to select on non-additive, time-varying idiosyncratic shocks. This does not mean that there is no selection on unobservables in the model. Indeed, I allow for rich differences in comparative advantage both across workers and through the life cycle, but idiosyncratic shocks pose difficulties. This is a departure from standard Roy models (e.g., Heckman and Honore (1990)) as well as previous papers in the structural labor literature. Unfortunately, when choice sets become very large, even fully specifying beliefs and solving the model make dealing with unobservable, time-varying multiplicative shocks difficult. This is because the solution to the value function would require very high dimensional integrals which even the best numerical methods cannot yet tackle. While incorporating this level of selection may be intractable, that does not validate ignoring it. To that end, one may think of the bias that arises from excluding this kind of selection. I work out the details in the technical appendix and I consider two special cases here.

First, suppose that occupations are symmetric in the sense that wages are the same for all and costs are symmetric—so only moving shocks determine movement. In this case, one can solve for the the main regression equation as,

$$\log \frac{P(o'|o;\sigma)}{P(o|o;\sigma)} = -C/\rho + \underbrace{\frac{w^2\sigma^2}{\rho^2} \times \frac{NP(o|o;0)-1}{N-1}}_{\text{Bias}}$$

where N is the number of occupations. The bias term is positive so that if $C > 0$, moving costs will be biased to 0. To the extent that one believes this approximation, it means that the costs will be underestimated and thus any estimated impact on workers can be viewed as a lower bound. The intuition for this result is that if there are many symmetric occupations, then the probability of getting both a positive moving cost shock and positive wage shock in an occupation other than the agent's current one is relatively high. Thus, workers will move with a higher probability than if there were just moving cost shocks—leading to an underestimation of moving costs.

In the asymmetric case, there is no closed form solution. However, by taking second order approximations around the variance of income shocks, one can demonstrate that the conditions for negative bias are quite stringent. In most empirically relevant cases, bias will tend to push cost estimates to zero, underestimating the impact to workers. In particular, the condition for the bias to not work in my favor is for low paying occupations to have substantially higher higher shock variances than high paying occupations, and while the correlation between variance and levels is negative, it is quite mild empirically.

5 Results

Here I present the major results from my estimation. First I discuss the model's fit—which is good on several metrics. Second, I discuss my findings on the costs of occupational switching and compare occupations and sectors. Finally, I discuss how costs vary with one's state and the importance of modeling unobserved comparative advantage.

5.1 Goodness of Fit

I asses the fit of model along two dimensions: in-sample fit and out-of-sample performance. First, I use the R^2 of my fitting procedure to determine the in-sample fit. The centered and uncentered[5] R^2 from my second stage estimation are .5353 and .9184, respectively. In interpreting these two numbers one should keep in mind that the model allows for 78 parameters to explain several million observations. Thus, unlike in a moment-matching procedure, the disparity between the number of parameters and number of targets is several orders of magnitude. Moreover, recall that the residual terms reflect worker

[5] I report the uncentered R^2 because a non-linear model lacks a natural interpretation for the R^2.

uncertainty and ought to have positive variance in the absence of perfect foresight. Because the variance of expectation errors is unknown, interpreting the R^2 is difficult. However, it's magnitude is large and reassures us that a relatively parsimonious model can explain a substantial amount of variation.

The second, and more telling, model assessment comes from looking at over-identifying restrictions. A natural choice of restrictions are the model's predictions for the distribution of occupational switches. The model is fit to discounted *relative* differences in switching probabilities, but the actual distribution of switches is a separate set of moments that can be used to benchmark performance. Figure 1 plots actual unconditional transition rates against the log of predicted unconditional transition rates. The dotted line is the 45 degree line while the black line is the best fit line. There are three takeaways. First, the overall fit to the EDF is very good, with a regression of actual on predicted having an R^2 .8557 and a coefficient of .747. Second, the fit is dramatically improved when weighting observations by their empirical likelihood—the R^2 increases and a slope coefficient of .889. Finally, the error is systematic with a regression coefficient of less than 1. This is a consequence of the multinomial logit approach which tends to equalize transition rates, leading me under predict the diagonal and have overly uniform transitions across occupations. This is particularly problematic when there are larger choice sets since the multinomial logit approach precludes 0 transitions, even though these are observed empirically. This could be remedied by arbitrarily limiting choice sets (i.e., setting mobility costs to infinity), but the model's fit is good without imposing these restrictions and allowing for this movement might be relevant in counterfactuals.

5.2 The Costs of Occupational Switching

Moving on to the costs of occupational switching, I have three major findings: first, switching costs can be large, on the order of ten years' income; second, switching costs are very heterogeneous across one's initial and target occupation, with the "cheapest" occupation being one third as costly to switch to as the most "expensive"; third, intrasectoral costs of occupational adjustment can be larger than intersectoral adjustment, thus ignoring the occupation margin several underestimates the impact of trade shocks on worker adjustment.

Turning to the first finding, figure 3 presents the observed histogram of costs conditional on switching. The distribution is right skewed with a median of 4.58, against a median income normalized to .425. This implies that the median costs of switching are on the order of 10 years median income. However, the median switcher is not the same as the median worker. To that end, figure 4 plots the histogram of switching costs over the income of each worker over the sample.[6] In the histogram the median is a cost of 11.1 years income and the overall distribution has more mass at the median and to its right. In interpreting these numbers, one must keep two things in mind. First, that switchers are overwhelmingly young so that the high ratio of costs to income partially reflects the life cycle. Second, these are costs exclusive of shocks. The shocks have an unconditional mean on the order of 10% of switching costs and the mean conditional on switching is higher still. Subtracting out these shocks implies that switching costs can still be as high as eight years' income. This explains the heavy diagonal in worker transitions, but also suggests that adjustment to unanticipated shocks can be costly.

Costs are not only large but vary substantially across one's initial occupations. This point is particularly important as understanding the variation across occupations is crucial to understanding the heterogeneous impacts of shocks. If all costs are the same, then all distributional consequences can be summarized by changes in relative wages, transition costs only act to slow down adjustment to a new equilibrium. On the other hand, if costs are disparate across occupation then they represent a source of distributional consequences beyond mere wages. They also imply that adjustment times may actually vary across groups and depend on which occupations are more or less affected by shocks. Figure 5 plots the density of the mean cost of moving from an occupation. That is to say, for a fixed source occupation, o, a unit of observation is $\frac{1}{|\mathcal{O}|}\sum_{o'} C(o, o')$. The density is unweighted by the composition of workers to avoid conflation with equilibrium outcomes. Instead, this figure reflects the rough level of costs that a worker in some occupation will face when deciding to move. It's clear from the figure that even with only 38 occupations, there is substantial heterogeneity in the mean cost of moving out of an occupation with the cheapest and the most expensive differing by a factor of 2.11.

Parallel to the fact that some occupations can be costly to enter, some occupations are costly to *exit*.

[6] I actually use the mean over the sample period. Using the mean serves two purposes. It helps reduce some noise in the income measure as periods of switching are associated with lower incomes (because of less time working) and it alleviates (partially) the problem that switchers tend to be younger and thus have lower incomes.

This has serious ramifications for understanding trade shocks. If trade shocks affect those occupations that are costliest to leave, then workers affected by shocks may opt to exit the labor market, becoming displaced, or simply accept permanently lower wages. Figure 6 plots the full heat map of costs from source to target occupations. These occupations are sorted by mean cost of exiting the occupation. Each row is the distribution of costs of moving from the source to the target occupation.This chart makes clear that some occupations are attractive in that exiting that occupation is very low cost, while other occupations can be costly to leave. Moreover, these costs *do* seem to be higher for those occupations more susceptible to trade shocks. Taking the average across all occupations within a sector, the costliest intersectoral transitions are into health and education, and these costs are highest for manufacturing occupations. Workers in manufacturing occupations also face high costs of entering services and FIRE. Most problematic, manufacturing occupations also face substantial within-sector adjustment costs. As manufacturing occupations are the most vulnerable to trade shocks, my results imply that trade shocks may result in particularly burdensome worker adjustments.

Finally, many models have focused on sectoral movement in order to model workers' adjustment to trade shocks. While I too have highlighted the importance of intersectoral movement, I have also stressed the substantial costs of changing occupations within a sector. That is to say, even if workers in manufacturing stay inside the manufacturing sector by, for example, moving from machine work to truck driving, this adjustment can be costly and painful. My model offers a natural way to tease apart costs across the sectoral and occupational dimension. And my results demonstrate that occupational movement is at least, if not *more*, important than movement across sectors. Figure 7 plots the densities of costs for workers switching sectors but not occupations, workers switching occupations but not sectors, and workers switching both. Costs for workers switching sectors is very tight while costs for workers switching occupations is very spread out. Table 5 summarizes the disparity in both magnitudes and spread. Intrasectoral occupation movement is associated with median costs that are 28% larger than pure sectoral movement. At the level of aggregation in the model, pure intrasectoral movement accounts for 42% transitions. These facts together imply that understanding worker adjustment requires tackling the occupational dimension.

5.3 The Role of the Worker's State

Tying to the introductory discussion of the life cycle, switching occupations becomes more costly with age, and this influences older worker's choice of occupation. Table 2 displays the estimates of the inverse productivity function. From the first two rows, a 35 year old should face 7% higher moving costs than a 25 year old. With the magnitudes under discussion, this can amount to nearly one additional year's income. This increased cost manifests itself in the decisions of actual workers. One can see this in table 3. Despite a large difference in costs between young and old workers *conditional on the same transition*, the observed difference between the youngest and oldest age group is only 7% at the median and is negative when comparing the second oldest group to the oldest. The reason is because older workers are more likely to choose occupations near to them in task space, lowering their actual moving costs. This highlights the importance of focusing on the composition of workers' ages when thinking about shocks.

Workers' skill and unobserved type actually matter as much as age. Looking again at the point estimates in table 2, the data selects a high type and a low type within each skill level. This highlights the importance of controlling for unobservable heterogeneity—aside from patterns in comparative advantage, the differences in moving cost efficiency imply that frequently moving workers may actually not be as adversely affected by a shock as workers who move little. And in fact, these differences are actually very large. For example, the mobility costs for type 2 medium skilled workers are 23% higher than for type 1 high skilled workers. In general, high skilled workers face lower mobility costs for the same occupational transitions. Perhaps surprisingly, middle skilled workers face the highest moving costs. However, as in the case of age groups, workers prefer closer occupations over high mobility costs. Table 4 summarizes the distribution of costs by group. Once again, the actual variation in median costs across groups pales in comparison to the cost differential conditional on the same transitions.

Finally, I briefly discuss my findings regarding comparative advantage and their importance in understanding worker transitions. I allowed for two unobserved types for each skill group—thus I have 6 types total. The first stage estimation, which required calculating and maximizing an empirical likelihood function, strongly rejects one type per skill level. The model selects a high type and a low type within each skill group, where high and low refer to the mean parameter value across occupations. Moreover, the differences across types is substantial. Figure 2 plots the comparative advantage of workers relative to high type of the high skilled workers. To read the graph, the x-axis reflects the comparative advantage of the high type and the y-axis is the comparative advantage of the low type. The unit of observation is an occupation. So, for example, the southernmost circle (corresponding to Teachers), says that the among low skilled workers, the high type is about ten percent worse than the high-type, high-skilled worker, while the low type worker is 250% worse. There are two major takeaways from the graph. First, on average, there is absolute advantage across types and skill groups—the majority of points are negative (meaning the higher skilled workers are better nearly all occupations). Second, comparative advantage is a potent force both across skills and types—this is visible from the variance present in the chart. If only absolute advantage were present then all the circles would clump at certain points. But quite to the contrary, there is substantial heterogeneity and often times one type's comparative advantage in a certain occupation overcomes their absolute disadvantage. For example, low-skilled, low-type workers are on average ten percent less productive than high-skill, high-type workers but are actually 17% more productive at retail and sales work. This may reflect that when initially deciding their skill level (a choice I do not model), workers with strong retail skills chose not to pursue higher education.

6 Counterfactuals

In this section I summarize my procedure for performing counterfactuals and perform several counterfactual exercises. To perform counterfactuals I need to take a stand on labor demand. To do so I specify a model of industry spillovers that allow for substantial heterogeneity in the occupational response to trade shocks. After outlining the model I describe how I calibrate key parameters of the demand side. Finally, I perform three counterfactuals. First, I simulate the Danish economy as if foreign prices were constant at their 2000 levels — before China's entry to the WTO and the inclusion of Eastern European countries in the EU. Second, I simulate the effects of a direct subsidy to switching costs for workers in occupations affected by trade shocks. In the third counterfactual I consider how early retirement programs can help smooth consumption losses to affected and displaced workers.

6.1 Closing the Model

In this subsection, I describe consumer preferences, the production side of the economy, and finally the model's equilibrium. In order to be consistent with the dynamic model in which workers maximize income, I propose a simple homothetic demand structure for consumers. The production side features the aforementioned industry spillovers.

6.1.1 Consumer's Preferences and Final Demand

I assume that workers life for a finite number of periods, $1, ..., T$. In each period they choose an occupation (potentially non-employment) and supply their time inelastically. Moreover, I abstract from savings decisions and assume that agents consume their entire incomes in each period. Thus workers' consumption decisions are static and decoupled from their working decision. In this section I specify only the consumption decision of workers and describe the occupational choice decision below. The consumers' utility function is a three-tiered utility function as in Broda and Weinstein (2006). The top tier is a Cobb-Douglas aggregator over industry-level outputs:

$$U = \prod_{i \in \mathcal{I}} C_i^{\alpha_i}$$

where $i \in \mathcal{I}$ indexes industries or goods and C is consumption of industrial aggregate i. The second tier is an Armington aggregator over domestic and foreign varieties:

$$C_i = \left(\left(C_i^D\right)^{\rho_i} + \left(C_i^F\right)^{\rho_i}\right)^{1/\rho_i}$$

where $\rho_i \in (0, 1)$ is an industry specific parameter and D and F refer to Denmark and Foreign respectively. Finally, the third tier is a CES aggregator over foreign varieties to construct the foreign aggregate.

Let $\sigma_i = \frac{1}{1-\rho_i}$ be the industry-specific elasticity of substitution. In this case, for a given level of expenditure in industry i, sE_i, the share of expenditure on domestic varieties will be given by,

$$E_i^D = E_i \frac{\left(P_i^D\right)^{1-\sigma}}{\left(P_i^D\right)^{1-\sigma} + \left(P_i^F\right)^{1-\sigma}}$$

Consumption is handled in a very simple way to focus on the richness of the labor market. To that end, I assume that export demand is exogenous and given by, $A_i^F \left(P_i^D\right)^{-\sigma}$where A_i is a demand shifter. Putting these pieces together yields the demand curve for domestic final production:

$$E_i^{D,final} = W\alpha_i \times \frac{\left(P_i^D\right)^{1-\sigma_i}}{\left(P_i^D\right)^{1-\sigma_i} + \left(P_i^F\right)^{1-\sigma_i}} + A_i^F \left(P_i^D\right)^{-\sigma_i}$$

where W is aggregate income. This is *not* total demand for good i. This is because my model features an input-output structure which I describe in the next section before returning to the workers' dynamic work decision.

6.1.2 Labor Demand: Representative Firm's Problem

In my model, the key to heterogeneity in the labor demand curve response to trade shocks is inter-industry linkages. By setting up production in this way, a system of industry-level Cobb-Douglas production functions give rise to complex and varied substitution patterns across occupations. In standard models, the focus is on broad sectors (e.g., manufacturing and services) or on aggregate output. I focus on highly disaggregated industries. This allows me to capture very rich aggregate substitution patterns documented elsewhere, such as capital-skill complementarity, while focusing on a very tractable underlying model. In particular, I posit "roundabout" production as in Foerster et al. (2011):

$$Y_i = z_i K^{\beta_i^K} \prod_{o \in \mathcal{O}} H_o^{\beta_{oi}^H} \prod_{j \in \mathcal{I}} M_i^{\beta_{ij}^M}$$

where K is capital, H is human capital and M refers to an industry aggregate used as an intermediate . Here, i and j index industries, and o indexes occupations. I assume that $\sum \beta = 1$ (but β can be 0 for some occupations and industries). As before I assume that M_i is an aggregator across domestic and foreign goods and the foreign good itself is an aggregator across foreign varieties. I assume the same elasticity of substitution is used by all industries and consumers. I assume a perfectly competitive representative firm in each industry. Moreover, I follow the common assumption in this literature and assume that world markets determine a perfectly elastic price of capital, r_i. This characterization leads to the following demand for industry output:

$$E_i^D = \left(\alpha_i W + \sum_{j \in \mathcal{I}} \beta_{kj}^M R_j\right) \times \frac{\left(P_i^D\right)^{1-\sigma_i}}{\left(P_i^D\right)^{1-\sigma_i} + \left(P_i^F\right)^{1-\sigma_i}} + A_i^F \left(P_i^D\right)^{-\sigma_i}$$

The strength of this setup is that, despite being simple, it allows for rich interactions in the economy through the input-output matrix, $\mathbf{B^M}$. Analyzing the market demand for occupations illustrates this flexibility. First, from the Cobb-Douglas system, one can derive the market expenditure on occupation o:

$$w_o H_o = \sum_{i \in \mathcal{I}} \beta_{oi}^H E_i^D$$

where E_i^D is domestic expenditure in industry i. Now consider a change in the price of foreign good i and hold all other prices (including wages) fixed. The, by substituting in the expression for domestic expenditure and totally differentiating with respect to P_k^F one has:

$$w_o dH_o = \beta_{ok}^H \underbrace{(\sigma - 1) \times \left[\left(\alpha_i^D W + \sum_{j \in \mathcal{I}} \beta_{kj}^M R_j\right) \frac{\left(P_k^D\right)^{1-\sigma}}{P_k} \times \frac{\left(P_k^F\right)^{-\sigma}}{P_k}\right] dP_k^F}_{\text{Import Competition Effect}} + \underbrace{\sum_{i \in \mathcal{I}} \beta_{oi}^H \left(\alpha_i^D dW + \sum_{j \in \mathcal{I}} \beta_{ij}^M dR_j\right) \frac{\left(P_i^D\right)^{1-\sigma}}{\left(P_i^D\right)^{1-\sigma} + \left(P_i^F\right)^{1-\sigma}}}_{\text{Input Price Effect}}$$

where the first effect maps out the impact of import competition on demand for industry k while the second effect maps out how various industries substitute from labor to other inputs as a result of changes in the relative price of inputs. Given that $\sigma > 1$, the direct effect of a drop in the price of k from F will be to lower demand for occupation o. However, the change in P_k^F changes the revenues in other industries who likewise adjust their demand for o. Thus, even the partial equilibrium effect of foreign prices on occupational demand can be arbitrary—varying across occupations and potentially dependent on which industry's prices change.

6.2 Equilibrium

Despite the complex economy-level substitution patterns between foreign supply and labor demand, this set up gives rise to a very simple equilibrium characterization. An equilibrium is defined by a set of domestic prices, $\{P_i^D\}_{i\in\mathcal{I}}$, wages $\{w_o\}_{o\in\mathcal{O}}$, labor stocks, $\{H_o\}_{o\in\mathcal{O}}$ and revenues, $\{R_i\}_{i\in\mathcal{I}}$ such that:

1. Representative firms choose intermediates, labor and capital optimally
2. Workers act optimally
3. Goods Market Clearing (for each $i \in \mathcal{I}$):

$$\underbrace{R_i}_{\text{Revenue}} = \underbrace{\left(\alpha_i^D W + \sum_{j\in\mathcal{I}} \beta_{ij}^M R_j\right)}_{\text{Domestic Exp. in } i} \underbrace{\frac{\left(P_i^D\right)^{1-\sigma}}{\left(P_i^D\right)^{1-\sigma} + \left(P_i^F\right)^{1-\sigma}}}_{\text{Domestic Share}} + \underbrace{A_i^F \left(P_i^D\right)^{1-\sigma}}_{\text{Foreign Exp.}}$$

4. Labor Market Clearing (for each $o \in \mathcal{O}$):

$$w_o \times \underbrace{\left(\sum_{\{n:o(n)=o\}} h_{on}\right)}_{\text{Labor Supply in } o} = \underbrace{\sum_{i\in\mathcal{I}} \beta_{oi}^H R_i}_{\text{Labor Demand}}$$

5. Balanced Trade:

$$\underbrace{\sum_{i\in\mathcal{I}} A_i^F \left(P_i^D\right)^{1-\sigma_j}}_{\text{Exports}} = \underbrace{\sum_i \left[\left(\alpha_i^D W + \sum_{j\in\mathcal{I}} \beta_{ij}^M R_j\right) \frac{\left(P_i^F\right)^{1-\sigma}}{\left(P_i^D\right)^{1-\sigma} + \left(P_i^F\right)^{1-\sigma}}\right]}_{\text{Imports}}$$

6.3 Demand Side Calibration

6.3.1 Procedure

Calibration proceeds in several steps. I discuss the details of the Danish IO matrices and how I handle changes over time in the Data Appendix, but I outline the procedure here. First, if inputs are flexible then production function coefficients on intermediates, capital and total labor can be estimated from expenditure shares and the IO matrix published in the Danish national accounts.

To construct the parameters for occupational expenditures I first calculate a labor coefficient from the Danish IO tables as above.[7] Then I use relative wage bills within industries to calculate β_{io}^{H}. Finally, I calculate β_i^K using gross surplus in the IO tables. This actually completes the estimation of the production parameters.

To estimate foreign prices, I use Danish customs data and the procedure of Broda and Weinstein (2006). The details are in the data appendix, but the method allows one to construct a CES price index for imported industry aggregates while taking explicit account of the fact that different countries produce goods of different quality.

Finally, to calibrate domestic prices notice that relative expenditure shares are a sufficient statistic for relative prices:

$$\frac{E_i^D}{E_i^F} = \left(\frac{P_i^F}{P_i^D}\right)^{1-\sigma}$$

With domestic prices in hand I can calculate the export shifters.

6.3.2 Calibration Results

Trade shocks enter this model through two channels, both of which are relative price effects. First, trade can change the relative demand for domestic tradable goods relative to non-tradable goods and services. Second, trade can change the composition of demand for domestic tradable goods, pushing workers from across narrow industries within manufacturing. Thus, for trade shocks to play a role in inducing occupational reallocation, we need to check to what extent demand for tradable goods has changed.

Figure 8 displays the change in log points from 1995 to 2005 in the ratio of domestic output to foreign imports. From the bar graph it's clear that manufacturing is on the decline and being replaced by imports. Also, there is substantial variation across industries in movement from domestic production to imports. For example, relative to imports, the tobacco and petroleum industries have actually grown while apparel has decreased by more than a full log point. This measure suggests ample room for reallocation of workers across and within sectors.

[7] The reason I do not use wage bills for the aggregate components is because of missing and imputed data. An implicit assumption is that missing and imputed data is random.

Changes in domestic production may conflate changes in trade conditions with domestic production shocks. To that end, figure 9 shows the log change in the foreign price index of different industries. The change in prices is less disperse than the change in domestic production shares. Nevertheless, the same two facts stand out. First, that foreign prices (relative to the Danish CPI) fell from 1995 to 2005 in nearly all industries. Second, this change in prices is not uniform, making room for trade to alter the composition of workers across industries.

6.4 Counterfactuals

In progress...

7 Appendices

7.1 Appendix: Tables

Table 1: Examples of Tasks

Com. 1: Communicative Activities	**Com. 2: Monitoring**
Written Expression	Far Vision
Written Comprehension	Operation Monitoring
Writing	Perceptual Speed

Com. 3: Facetime Tasks	**Com. 8: Routine Maintenance**
Performing for or Working Directly with the Public	Installation
Assisting and Caring for Others	Repairing
Resolving Conflicts and Negotiating with Others	Equipment Maintenance

Table 2: Inverse Switching Productivity Parameters

$f(\omega) = \exp\left(\beta_a^f * a + \beta_{a^2}^f a^2 + \sum_i \beta_i^f \times D_i\right)$	
Age Params.	
β_a^f	.0158
$\beta_{a^2}^f$	1.501×10^{-4}
Type Params.	
β_1^f	-0.089
β_2^f	0.076
β_3^f	-0.021
β_4^f	0.105
β_5^f	-0.106
β_6^f	0

Table 3: Mobility Costs by Age Group

Age	Mean	Q_{25}	Q_{50}	Q_{75}
25-29	4.89	3.85	4.54	5.70
30-39	5.03	3.97	4.69	5.86
40-49	5.25	4.15	4.91	6.14
50-59	5.27	4.15	4.88	6.18

Table 4: Mobility Costs by Skill and Type

Skill Type	Mean	Q_{25}	Q_{50}	Q_{75}
Low T1	4.50	3.58	4.16	5.34
Low T2	5.21	4.17	4.83	6.01
Med T1	4.65	3.77	4.30	5.36
Med T2	5.40	4.28	5.08	6.24
High T1	4.40	3.41	3.99	5.17
High T2	4.90	3.80	4.61	5.74

Table 5: Mobility Costs by Transition Type

Transition Type	Mean	Q_{25}	Q_{50}	Q_{75}
Sector Only	4.10	3.81	4.09	4.38
Occ. Only	5.32	4.20	5.26	6.15
Both	5.72	4.61	5.67	6.68

7.2 Appendix: Figures

Figure 1: Model Fit: Predicted Versus Actual Transitions

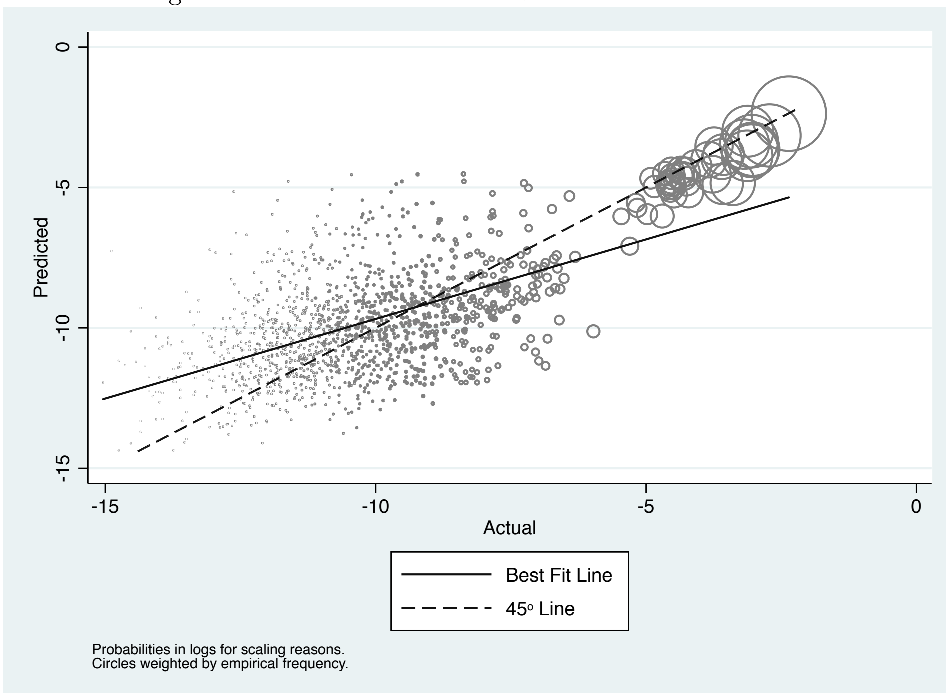


Figure 2: Comparative Advantage

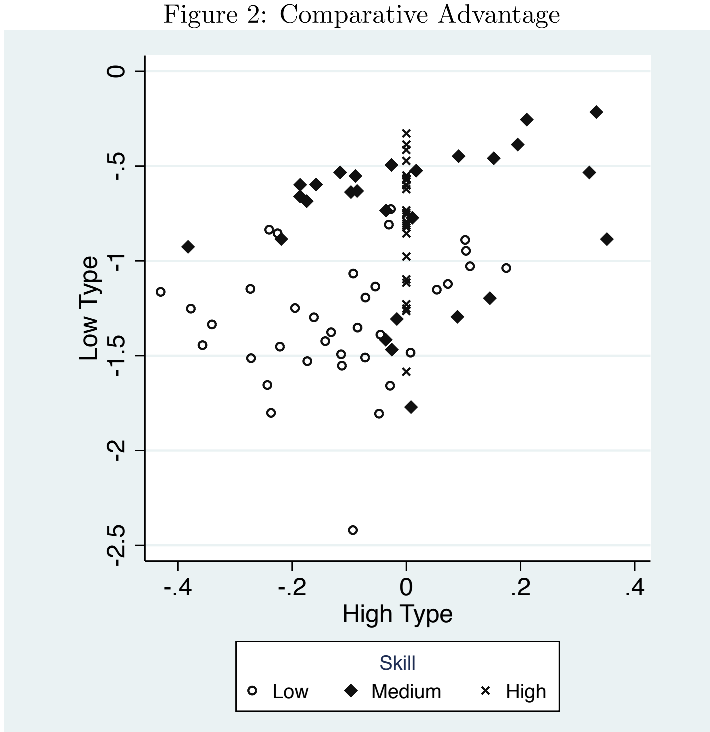


Figure 3: Histogram of Switching Costs

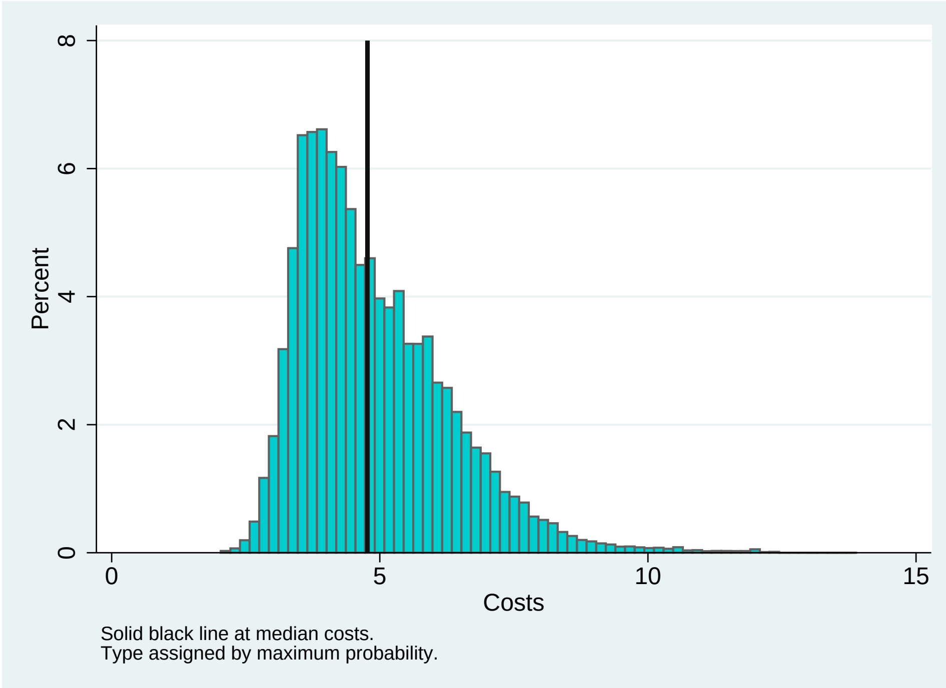


Figure 4: Histogram of Switching Costs Relative to Income

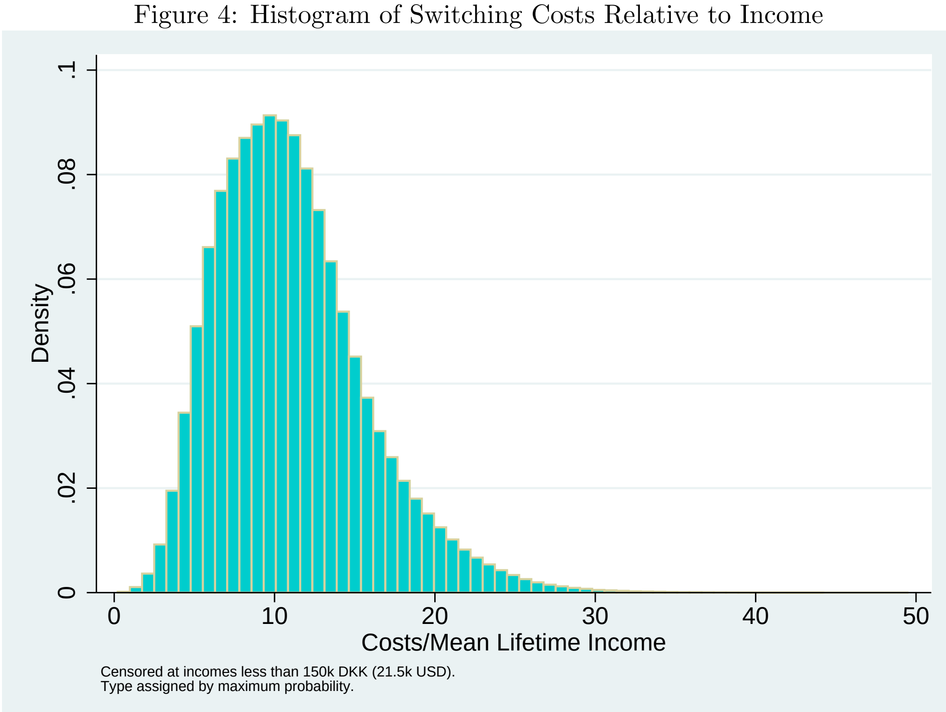


Figure 5: Density of Mean Costs Out of Source Occupation

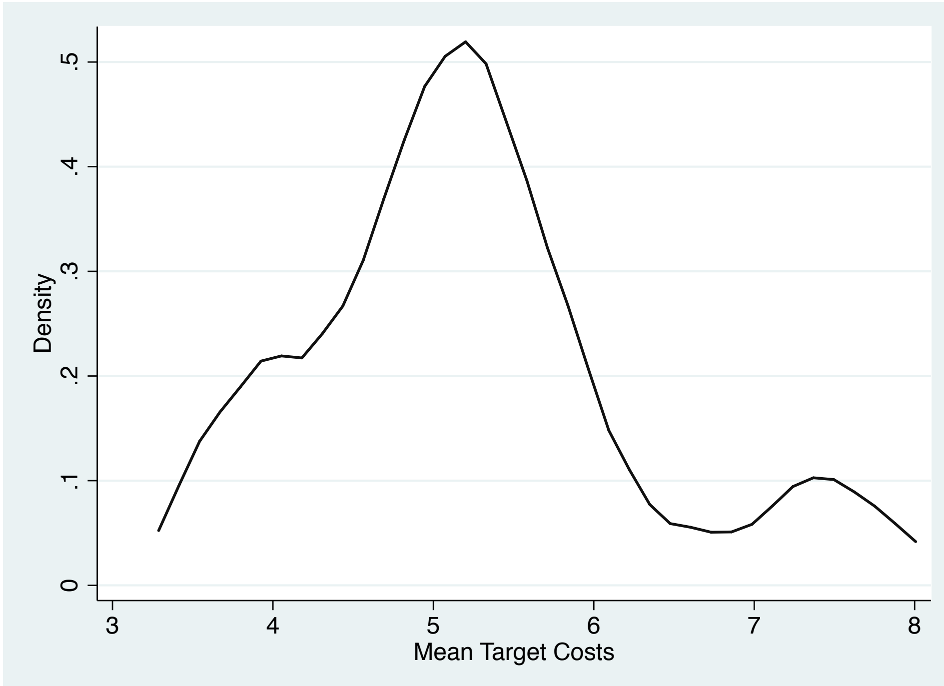


Figure 6: Costs of Moving from Source Occupation to Target

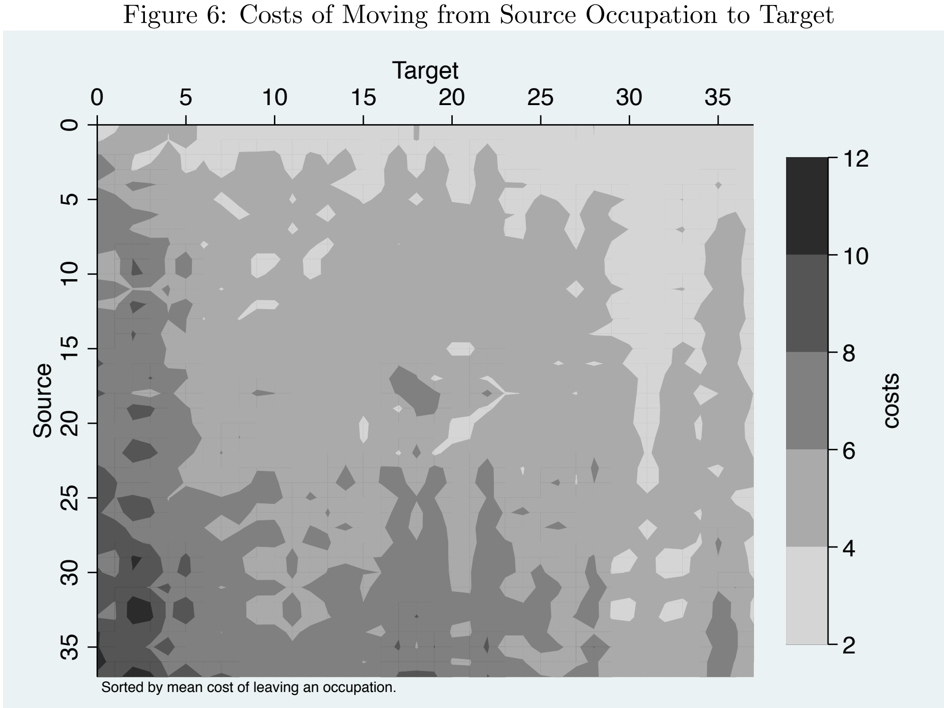


Figure 7: Density of Switching Costs by Type of Transition

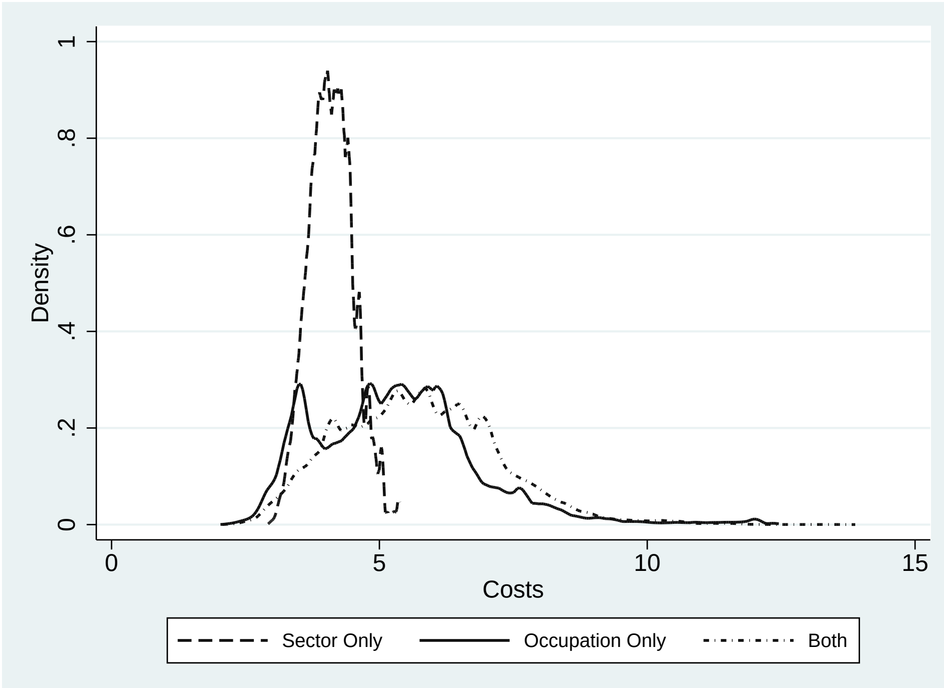


Figure 8: Changes in Domestic Production Relative to Imports

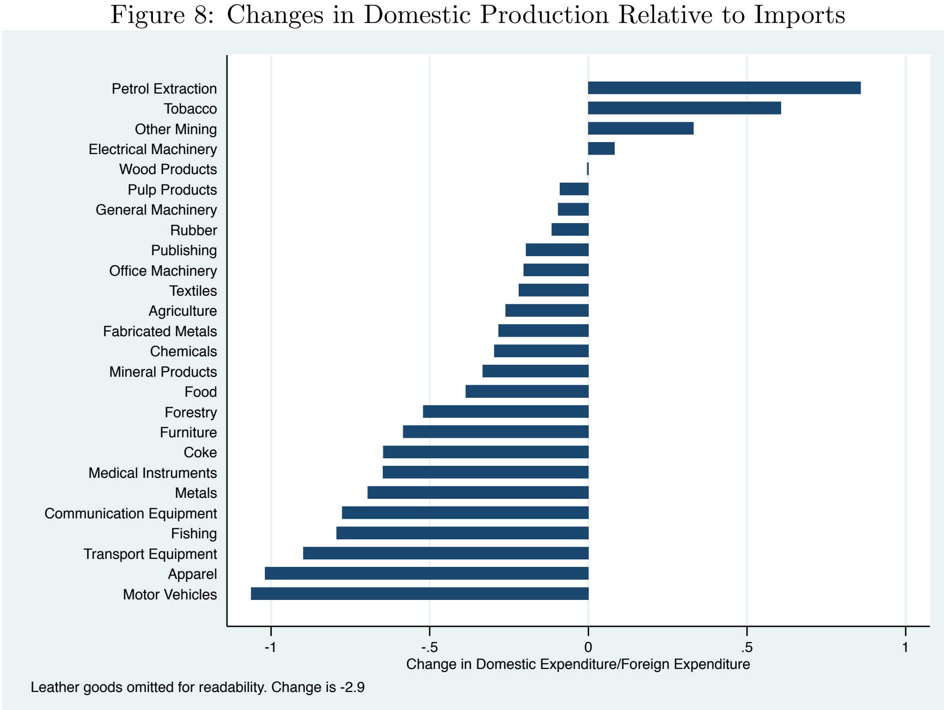
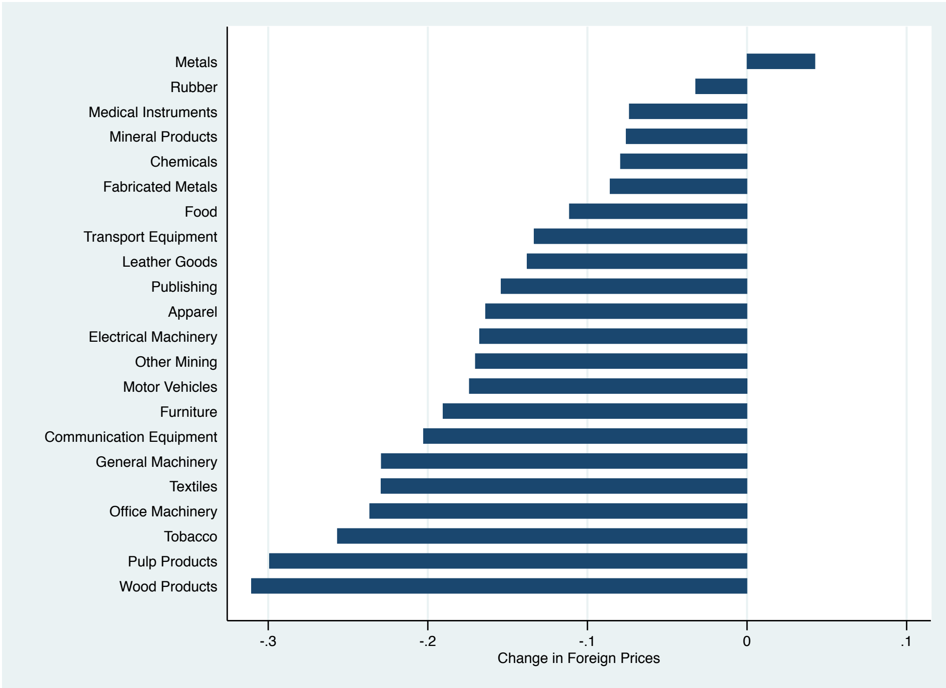


Figure 9: Changes in Foreign Prices



References

Abowd, John, Francis Kramarz, and David Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–334.

Acemoglu, Daron and David H. Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Ashenfelter O. and Card D., eds., *Handbook of Labor Economics*, 1 ed., Vol. 4B, Elsevier, 2011, chapter 12, pp. 1043–1171.

Altonji, Joseph G., Anthony A. Smith Jr., and Ivan Vidangos, "Modeling Earnings Dynamics," *Econometrica*, 07 2013, *81* (4), 1395–1454.

Altug, Sumru and Robert A Miller, "The Effect of Work Experience on Female Wages and Labour Supply," *Review of Economic Studies*, January 1998, *65* (1), 45–85.

Arcidiacono, Peter and Robert A. Miller, "CCP Estimation of Dynamic Models with Unobserved Heterogeneity," *Econometrica*, November 2011, *79* (6), 1823–1868.

Artuc, Erhan and John McLaren, "Trade Policy and Wage Inequality: A Structural Analysis with Occupational and Sectoral Mobility," NBER Working Papers 18503, National Bureau of Economic Research, Inc Nov 2012.

_ , Shubham Chaudhuri, and John McLaren, "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, June 2010, *100* (3), 1008–45.

Autor, David H., David Dorn, Gordon H. Hanson, and Jae Song, "Trade Adjustment: Worker-Level Evidence," *The Quarterly Journal of Economics*, 2014, *129* (4), 1799–1860.

Bai, Jushan and Serena Ng, "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, January 2002, *70* (1), 191–221.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–90.

Broda, Christian and David E. Weinstein, "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, May 2006, *121* (2), 541–585.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg, "The Anatomy of French Production Hierarchies," CEPR Discussion Papers 9073, C.E.P.R. Discussion Papers 2012.

_ , Maximiliano Dvorkin, and Fernando Parro, "The Impact of Trade on Labor Market Dynamics," Working Papers 2015-9, Federal Reserve Bank of St. Louis 2015.

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

Crawford, Gregory S. and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73* (4), 1137–1173.

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

Farber, Henry S and Robert Gibbons, "Learning and Wage Dynamics," *The Quarterly Journal of Economics*, 1996, *111* (4), 1007–47.

Foerster, Andrew, Pierre Daniel Sarte, and Mark Watson, "Sectoral versus Aggregate Shocks: A Structural Factor Analysis of Industrial Production," *Journal of Political Economy*, 2011, *119* (1), 1 – 38.

Grossman, Gene M. and Esteban Rossi-Hansberg, "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, 2008, *98* (5), 1978–97.

Heckman, James and Burton Singer, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 1984, *52* (2), 271–320.

Heckman, James J and Bo E Honore, "The Empirical Content of the Roy Model," *Econometrica*, September 1990, *58* (5), 1121–49.

Hotz, V. Joseph and Robert A. Miller, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, July 1993, *60* (3), 497–529.

Jacobson, Louis S, Robert LaLonde, and Daniel Sullivan, "Earnings Losses of Displaced Workers," *American Economic Review*, 1993, *83* (4), 685–709.

Kasahara, Hiroyuki and Katsumi Shimotsu, "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices," *Econometrica*, 01 2009, *77* (1), 135–175.

Keane, Michael and Kenneth Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, 1997, *105* (3), 473–522.

_ , Petra Todd, and Kenneth Wolpin, "The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications," in Ashenfelter O. and Card D., eds., *Handbook of Labor Economics*, 1 ed., Vol. 4A, Elsevier, 2011, chapter 04, pp. 331–461.

Lee, Donghoon and Kenneth I. Wolpin, "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica*, 01 2006, *74* (1), 1–46.

Morales, Eduardo, Gloria Sheu, and Andrs Zahler, "Gravity and Extended Gravity: Using Moment Inequalities to Estimate a Model of Export Entry," NBER Working Papers 19916, National Bureau of Economic Research, Inc February 2014.

Rust, John, "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, September 1987, *55* (5), 999–1033.

Scott, Paul, "Dynamic Discrete Choice Estimation of Agricultural Land Use," *Mimeo*, 2014.

Stock, James and Mark Watson, "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business & Economic Statistics*, 2002, *20* (2), 147–62.