Globalization and Labor Market Dynamics.*

John McLaren (University of Virginia).

August 15, 2016

Contents

1 A missing tool, and why it matters. 3

2 Adding dynamic labor adjustment to the trade-theory toolkit. 4
   2.1 Essential elements for a useful theory. . . . . . . . . . . . . . . . . . . 4
   2.2 A bare-bones model. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
   2.3 Embedding dynamics in a rich trade model. . . . . . . . . . . . . . . . . . 10
   2.4 Other approaches: Search frictions, firing costs, and generational change. . 12

3 Structural empirical approaches. 16
   3.1 First wave: No unobserved heterogeneity. . . . . . . . . . . . . . . . . . . 17
   3.2 Approaches with unobserved heterogeneity. . . . . . . . . . . . . . . . . . 26
   3.3 Observations. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32

4 Issues for further exploration. 33

Abstract

Historically, the trade research field has usually ignored dynamic adjustment of workers, but a recent wave of work has developed a rich set of theoretical and empirical tools to analyze this. Empirical approaches have ranged from reduced-form regressions to structural estimation of underlying parameters, which is necessary to get at welfare effects. A major distinction is between models that do and do not allow for unobserved heterogeneity across workers; these models are useful for different purposes. A consistent finding across methods and countries is that costs of switching

*Department of Economics, University of Virginia, P.O. Box 400182, Charlottesville, VA 22904-4182; jmclaren@virginia.edu.
sectors and occupations are high, and that both switching costs and option value are crucial in computing welfare effects of globalization for workers.
1 A missing tool, and why it matters.

If one accepts “Who benefits and who is helped by globalization, and by how much?” as the primal question of international trade research, then the centrality of adjustment costs for labor follows immediately. If restrictions on imported autos are lowered, the question of whether or not autoworkers are harmed, and how much, depends on how easily they can move to another industry; the question of whether or not workers in an auto-making town are harmed depends on how easily workers can move. Old-style trade theory, of course, was always very crude about this: Workers were assumed to be able to switch industries costlessly or not at all. The Heckscher-Ohlin model, the Ricardian model, and their cousins, all assumed that workers could switch industries without cost, and so if two industries are producing, they must offer the same wage to workers with the same ability. However, we know that this is not the way things work in the world. We have abundant evidence that there are frictions from switching industry,\(^1\) from moving across locations within a country,\(^2\) and from switching occupations.\(^3\) These all imply that whether one gains or loses from a trade shock may be determined by one’s industry, location, occupation, or similar factors.

Further, once we admit costs of switching into our trade theory, we need a dynamic model. The essence of switching industries becomes incurring a cost now in order to gain an advantage in future earnings – an investment. This is underlined by the abundant evidence that the labor-market response to trade shocks tends not to be quick. Menezes-Filho and Muendler (2011), studying Brazilian longitudinal data, show that workers separated from a job due to import competition tend to take years to find work in a new industry (often apparently spending long intervals in the informal sector in between). Autor, Dorn, Hanson and Song (2014), using longitudinal data on US workers, find effects of rising Chinese imports on affected blue-collar workers that linger more than a decade. Dix-Carneiro and Kovak

---

\(^1\)Revenge (1993) shows in US data that industry-level trade shocks show up in industry-level wage effects. Hakobyan and McLaren (2016) show similar effects from NAFTA. Developing-country evidence is abundant, for example, Currie and Harrison (1997) for Morocco.

\(^2\)Topalova shows that national trade shocks have differential local impacts in India, Kovak (2013) in Brazil, and Autor, Dorn and Hanson (2013) in the US. In each case, this implies moving costs that prevent workers from arbitraging away the local effect of shocks.

\(^3\)Ebenstein et al (2014) find differential effects of trade and offshoring shocks by occupation, indicating that they cannot be arbitraged away by easy occupation switching. Cortes and Gallipoli (2014) and Liu and Trefler (2011) also provide evidence on occupational-switching costs.
(2015) show that regional wage differentials caused by the differential local effect of the 1991 Brazilian trade liberalization were not eliminated or even narrowed even 20 years later (in fact, they widened). Hummels et al (2014) find negative effects on blue-collar wages in Danish firms that increase offshoring, which persist for 4 years.

In the real world dynamic adjustment is extremely important. Workers in a declining industry that is hit with accelerated imports due to a trade agreement do not pack up at once and switch to an expanding industry. Switching industries is costly and time-consuming, and for many workers it is never really worth it, so reallocation of labor occurs gradually and the fortunes of a worker in an afflicted industry can be very different from those of other workers. Dynamic adjustment must be included as part of the economist’s theoretical and empirical toolkit. In recent years an new wave of work has targeted these questions. We will review theoretical contributions and then empirical work, and sift through what has been learned and what new directions remain promising.

2 Adding dynamic labor adjustment to the trade-theory toolkit.

2.1 Essential elements for a useful theory.

In order to be useful in analyzing labor adjustment to a globalization shock, a theoretical model should have (at least) three features.

(i) *It should feature gross flows in excess of net flows, including churning of workers across jobs or locations in the steady state.* This is a robust feature of the data; for example, Artuç, Chaudhuri and McLaren (2010) document that in their data gross flows of workers across US sectors are about 10 times as large as net flows. Put differently, there are always large numbers of workers switching in opposite directions. This is one element lacking from earlier work, such as Dehejia (2002), based on convex-adjustment cost models of capital such as Mussa (1978).

(ii) *It should be a general-equilibrium model, so that we can analyze how wages across the economy respond to a globalization shock.* A drop in the price of an imported good will, in and of itself, raise the real wage of workers throughout the economy, and movements of
workers out of one sector will have an effect on wages in their origin sector and in their destination sector. A long tradition in reduced-form regressions assumes implicitly that worker outcomes in one industry are affected only by trade shocks in that industry (Revena (1992), for example), but that misses much of the story.

(iii) *It should feature forward-looking workers.* Many changes in trade and trade policy are anticipated. Workers read about negotiations of the North American Free Trade Agreement for years before it was signed and ratified. The end of the Multifiber Arrangement that had long protected rich-country textile and apparel industries was negotiated with a 10-year dismantling of the quotas, with most of the liberalization occurring in the last year. Workers in those industries knew what was coming a long time in advance. There is a lot of evidence that anticipation effects matter. Pierce and Schott (2016), for example, have shown evidence that China joining the WTO caused a large labor-market response by eliminating the possibility of future trade sanctions that had always been anticipated with positive probability. Recently the Brexit vote in the United Kingdom was in effect an announcement of a significant future change in trade policy, and it exhibited a sharp, immediate response in the British economy. (See Artuç et al (2014, Section 6) for more examples.)

These features allow us to analyze the labor market’s response to a globalization shock as consisting of (a) an announcement effect; (b) an impact effect; (c) the transition to the new steady state; and (d) the new steady state. All are necessary for a full account.

### 2.2 A bare-bones model.

To make the basic theoretical issues clear, consider the simplest possible model, along the lines of Artuç, Chaudhuri and McLaren (2008, 2014).

A small open economy has two sectors, X and Y, each of which combines a sector-specific fixed factor with labor. (There are no occupations, and each industry produces in one location.) There is a mass of workers of measure $L$. Output in sector $i$ in period $t$ is given by:

$$q_i^t = Q^i(L_i^t),$$

where $L_i^t$ denotes labor used in sector $i$ in period $t$ and we suppress the fixed factor in each industry to simplify notation; $Q^i(\cdot)$ is increasing, continuously differentiable, and strictly
concave.

The domestic price of good $i$, $p^i$, is equal to the given world price plus a tariff (export subsidy) for a good that is imported (exported). For now, assume that these prices are constant. Letting $\tilde{w}_i^t$ be the nominal wage paid by sector $i$ at date $t$ and $p^i$ the domestic price of good $i$, the wage in sector $i$ at date $t$ is the marginal value product of labor:

$$\tilde{w}_i^t = p^i \frac{\partial Q^i(L_i^t)}{\partial L_i^t},$$

(1)

In effect, in each period in each sector at each moment there is a vertical supply curve for labor, whose intersection with the downward-sloping labor-demand curve determines the sectoral wage.

The heart of the model is the mobility of labor. A worker $\theta$ who is in industry $i$ at the end of period $t$ receives a non-pecuniary benefit $\varepsilon_{\theta,t}^i$. This can be thought of as enjoyment of the work or of living in the region where industry $i$ is located, for example. Assume that the $\varepsilon_{\theta,t}^i$ are independently and identically distributed across workers and sectors and over time, with cdf $F(\cdot)$, pdf $f(\cdot)$, full support on the real line, and mean zero. The cost to a worker $\theta$, who was in $i$ during period $t$, of moving from $i$ to $j \neq i$ at the end of $t$ is, then:

$$C + \varepsilon_{\theta,t}^i - \varepsilon_{\theta,t}^j,$$

where $C \geq 0$ is the deterministic component of mobility costs, which we will call the ‘common’ portion since it is common to all workers. The variable $\varepsilon_{\theta,t}^i - \varepsilon_{\theta,t}^j$ is the idiosyncratic component of moving costs, which can be negative as easily as it can be positive. For example, a worker may be bored with her current job and long for a change in career ($\varepsilon_{\theta,t}^i - \varepsilon_{\theta,t}^j < 0$), and a worker with a child in the final year of high school may have a non-pecuniary reason to stay in the current location rather than move, as might be necessary to change jobs ($\varepsilon_{\theta,t}^i - \varepsilon_{\theta,t}^j > 0$). Since what is important for workers’ decisions is the difference between $\varepsilon^i$ and $\varepsilon^j$, we can simplify notation by defining:

$$\mu_{\theta,t}^i = \varepsilon_{\theta,t}^i - \varepsilon_{\theta,t}^j,$$

for a worker currently in sector $i$. The distribution of $\mu_{\theta,t}^i$ is derived from $F(\cdot)$ and $f(\cdot)$, and
the cdf and pdf are denoted by $G(\cdot)$ and $g(\cdot)$ respectively.

The assumption that these costs are independent across time is not realistic, but it makes the model vastly more tractable, and this class of models has made a lot of empirical headway. The more advanced models with unobserved heterogeneity will be discussed in Section 3.2.

The fraction of workers in sector $i$ at time $t$ who choose to switch to sector $j$ is denoted by $m_{ij}^t$, and so the allocation of labor evolves as:

$$m_{ii}^t L_i^t + m_{ji}^t L_j^t = L_i^t + 1$$

(2)

The gross flows of labor are then $m_{ij}^t L_i^t + m_{ji}^t L_j^t$, and the net flows are $L_{t+1}^i - L_i^t$. Workers have perfect foresight. Each worker takes the sequence of current and future wages in both sectors as given, and chooses a sector in each period to maximize expected discounted utility. All workers doing so at the same time induce an allocation of labor across the two sectors over time, which induces a sequence of wages over time through (1).

All agents are risk neutral, have rational expectations and have a common discount factor $\beta < 1$. Further, all workers have identical and homothetic preferences, which allows us to identify a common cost-of-living index, which we can write as $\phi(p^X, p^Y)$. Thus, the real wage is $w_i^t \equiv \tilde{w}_i^t / \phi(p^X, p^Y)$. Dropping the $\theta$ subscript, let $u^i(L_t, \epsilon_t)$ denote the (maximized) value to a worker of being in $i$ given $L_t = (L^X_t, L^Y_t)$ and idiosyncratic shocks $\epsilon_t = (\epsilon^X_t, \epsilon^Y_t)$ realized by the worker. Then $v^i(L_t) \equiv E_\epsilon(u^i(L_t, \epsilon_t))$ gives the expected value of $u^i$ before idiosyncratic shocks are realized, but conditional on $(L_t)$.

Since the worker is optimizing, $u^i(L_t, \epsilon_t)$ can be written:

$$u^i(L_t, \epsilon_t) = w_i^t + \max\{\epsilon_i^t + \beta v^i(L_{t+1}), \epsilon_j^t - C + \beta v^j(L_{t+1})\}$$

$$= w_i^t + \beta v^i(L_{t+1}) + \epsilon_i^t + \max\{0, \bar{\mu}_i^t - \mu_i^t\},$$

where

$$\bar{\mu}_i^t = \beta[v^j(L_{t+1}) - v^i(L_{t+1})] - C,$$

(3)

and $i \neq j$. The expression $\bar{\mu}_i^t$ is the common value of the net benefit of moving from $i$ to $j$. 

7
Any worker will switch sectors if and only if \( \mu_i^t < \pi_i^t \), and so \( m_{ij}^t = G(\pi_i^t) \) Taking expectations yields:

\[
v^i(L_t) = w^i_t + \beta v^i(L_{t+1}) + \Omega(\pi_i^t),
\]

where \( \Omega(\mu) = E\mu \max\{0, \mu - \mu\} \) can be interpreted as the worker’s option value.

In other words, the value, \( v^i \), of being in \( i \) is the sum of: (i) the wage, \( w_i^t \), that is received; (ii) the value, \( \beta v^i(L_{t+1}) \), of staying on in \( i \); and (iii) the additional value, \( \Omega(\pi_i^t) \), of having the option to move.

From a given starting point \( (L_0^X, L_0^Y) \), an equilibrium is then a sequence of \( \pi_i^t \) values for \( i = X, Y \) such that if every worker follows the rule of switching from \( i \) if and only her realization of \( \mu_i^t \) is below \( \pi_i^t \), the allocation of labor over time will induce a sequence of real wages in the two sectors such that that switching rule will be optimal for each worker. It can be shown that any equilibrium maximizes the present discounted real value of revenue in the two sectors net of moving costs (meaning both the common values \( C \) and the realized \( \mu_i^t \) values for workers who move). This can be called the ‘social planner’s problem,’ although care must be taken to note that if tariffs are non-zero it is a distorted social planner’s problem. Further, any solution to that same optimization problem is an equilibrium. Since it can be shown that this social planner’s problem is well-behaved and concave, there is a unique solution. As a result, there is a unique equilibrium. These points are developed in Cameron, Chaudhuri, and McLaren (2007).

Various properties that are useful for thinking about globalization in the real world are derived in Artuç et al (2014). All of these can be proven from the social-planner’s problem.

(i) Steady-state properties. First, regardless of the initial allocation of labor, for any given output prices (hence for any given tariffs), there is a unique steady state. Importantly, that is in general not a steady state in which wages are equalized across sectors.

This point can be understood as follows. Suppose that the real wages in both sectors were expected to be constant over time, and the difference is \( \Delta w \equiv w^X - w^Y \). Working through (3) and (4), this would imply a constant value of \( \pi_i^t \) which is an increasing function of \( \Delta w \) and a constant value of \( \pi_i^t \) which is a decreasing function of \( \Delta w \). Since \( m_i^{XY} \) and \( m_i^{YX} \) are increasing functions of \( \pi_i^X \) and \( \pi_i^Y \) respectively, then from (2) the steady-state share
of workers in Sector X must be increasing in $\Delta w$. Put differently, the long-run elasticity of intersectoral labor supply with respect to intersectoral wage differentials is positive and finite. In a model with no mobility costs, it would be infinite (and there would never be wage differences across sectors), and in a model with infinite mobility costs it would be zero. And with the mobility costs, only if the steady-state size of the two sectors is equal can the wages be equalized.

(ii) Response to a globalization shock. Suppose that Sector Y is the import-competing sector and that the economy is initially in a steady state with a permanent tariff. Suddenly, at $t = 0$, the tariff is removed. The real domestic price of sector-Y output drops, so the demand curve for labor in sector Y falls. The real wage in Y falls abruptly (by less than the drop in the tariff, because the consumer price index $\phi(p^X, p^Y)$ also falls), and the real wage in X rises abruptly because the consumer price index falls. Flows of workers out of Y increase, and flows of workers out of X decrease, so that workers reallocate toward X. As this happens, the equilibrium moves down and the left along the sector-X labor-demand curve, pushing down the wage there; and the opposite happens in sector Y. Therefore, the wages in the liberalizing sector drop abruptly when the tariff is removed, and then creep up over time, while the real wage in the export sector jumps up abruptly, and creeps down over time. In the new steady state, the real wages could be higher or lower than they were initially, but what we know for sure is that $w^X - w^Y$ will be higher than it was at the start (remember, after all, that the intersectoral elasticity of labor supply is positive, so to maintain a larger share of the economy in the export sector, its relative wage must be higher than it was at the start).

Note that adjustment to the shock will be gradual, because of the idiosyncratic shocks. In each period there will be workers who wish to move from Y to X, but who at the moment have high moving costs, and so they wait for a better time.

(iii) Welfare. As noted at the outset, the main interest is in who gains and who loses, which means whether the present discounted value of utility for a worker in sector $i$ goes up or down at date 0 when the tariff is removed. Because of the time-path of future wages and because of the importance of option value (recall (4)), the fact that the X wages rise on impact and the Y wages fall on impact is insufficient to determine this. In fact, as a result of the globalization shock it is possible for (a) workers in both sectors to benefit; (b) workers
in both sectors to be harmed; or (c) workers in the export sector to benefit and workers in the import-competing sector to be harmed.\(^4\) The one property that is guaranteed is that the net benefit to the workers in the export sector, whatever its sign, is better than the net benefit to workers in the import-competing sector.

\((iv)\) Delay, and announcement effects. Suppose that the economy is in a steady state with a tariff expected to stay in place forever, and then at date 0 a credible announcement is made that the tariff will be eliminated at date \(T > 0\), and never brought back. In this case, the economy’s adjustment will begin at date 0. Workers in the protected sector who happen to have low moving costs at the moment, anticipating a big drop in wages at date \(T\), will move to the export sector. This anticipatory migration accelerates as \(T\) approaches. During this anticipatory period real wages in the protected sector creep upward – above their pre-announcement levels – and newly arriving workers in the export sector push wages down there – below their pre-announcement levels. In effect, workers stuck in the protected sector during the period from 0 to \(T\) receive an anticipation rent, as their wages are temporarily elevated. Because of these processes, delaying liberalization instead of springing it as a shock-therapy surprise at date 0 raises the net lifetime utility benefit of protected-sector workers from the liberalization and lowers the net benefit of export-sector workers. A Unanimity Theorem emerges: If \(T\) is large enough, workers in both sectors will agree on the sign of the benefit, and will either both be for the liberalization or both against it.\(^5\)

2.3 Embedding dynamics in a rich trade model.

Models of the class described above are weak on the production side and on the trade side. They are based on a small open economy with a finite number of sectors, each producing

\(^4\)Outcome (a) is more likely, the higher the elasticity of substitution between labor and capital in the export sector relative to the import-competing sector, and vice versa for outcome (b). Intuitively, eliminating the tariff raises the demand for workers in the export sector and lowers it in the import-competing sector. The net effect on workers depends on which is stronger, and more substitutability between capital and labor in a sector allows for a bigger shift in labor demand.

\(^5\)Conditions analogous to the previous footnote determine which of these two cases applies. Note that a similar result is developed in an earlier paper by Dehejia (2002), who studies a labor-abundant Heckscher-Ohlin economy with two industries in which capital can be moved costlessly between industries but labor has convex industry-switching costs. A tariff initially protects the import-competing industry. In the absence of switching costs, all workers would be in favor of trade liberalization by Stolper-Samuelson logic, but with the switching costs workers initially in the import-competing industry oppose it. It is shown that in this case a sufficiently gradual removal of the tariff will result in all workers favoring the reform.
a single homogenous good, and there is no intra-industry trade or trade in intermediate inputs. There is a reason for that; dynamics creates complication, and keeping track of multiple countries and multiple industries, each with many varieties of good and a rich input-output structure – such as in the quantitative trade literature that has grown out of Eaton and Kortum (2002) (EK) – precludes any attention paid to dynamic adjustment. Caliendo, Dvorkin and Parro (2015), however, do what one might have thought impossible: They combine a serious model of labor-market dynamics with a very rich EK-type model in a way that preserves the tractability of the EK approach and allows for quantitative evaluation. Caliendo et al create a model with 38 countries, 22 sectors, and 50 states within the US. Workers are homogenous but face common and idiosyncratic costs of switching sectors and of moving across states. Each sector produces a continuum of goods with the productivity of each good’s production in each state and country given by the realization of a Frechet distribution as in EK, so that each local economy both exports and imports some of the goods in each industry.

One trick that is used heavily in quantitative trade models is the ‘exact hat algebra’ first used in Dekle, Eaton and Kortum (2008), which allows one to evaluate changes in a trade equilibrium in an EK-type model (such as, for example, due to a trade liberalization) without having to compute the whole equilibrium. Now, in Caliendo et al (2015), a version of the ‘exact hat algebra’ is developed that allows one to evaluate changes in the time-path of adjustment in an analogous way. In this setting, such a trick is essential to getting anywhere at all with the problem: With multiple countries and multiple industries as well as 50 states within the US, solving the dynamic equilibrium (in which each worker must form expectations about the future path of wages in each industry, in order to make a rational choice about switching industries or switching states at each date) would impose enormous computational demands. The extension of the ‘exact hat algebra’ to the dynamic setting makes it actually feasible.

The model is calibrated, based on estimated parameters from a variety of sources, and the model is used to simulate an increase in productivity in Chinese manufacturing, with the technology shock across sectors chosen to produce a surge in manufactured exports to the US similar to what is observed in the data. Over ten years, this causes a large movement of workers out of manufacturing, and the simulation can explain about half of the drop in
US manufacturing employment over this period. However, in every sector-state cell, the expected lifetime utility of a worker goes up at the date of the revelation of the China shock, relative to the pre-shock steady state. These gains to workers flow from three sources. One is the importance of option value, as mentioned above (recall equation (4)); even for a worker whose wages fall, the possibility of being able to switch in the near future to another industry whose wages have increased adds to the expected lifetime welfare of the worker. A second is the benefit of low-cost imports to workers as consumers. A third, very important but more subtle source of benefit is the improved productivity to each US industry due to the availability of low-cost imported inputs from China. These are all quantified in the paper. These optimistic results could be taken as a challenge to the interpretation of the well-known regression results of Autor, Dorn and Hanson (2013) and Autor, Dorn Hanson and Song (2014) that US workers in locations dependent on China-vulnerable industries were harmed by rising Chinese exports, and can be taken as a reminder of the importance of option value. We will return to this question in Section 3.1 (with a note of caution).

2.4 Other approaches: Search frictions, firing costs, and generational change.

The models sketched above generate gross flows in excess of net flows through idiosyncratic preference shocks. An equally appealing source of gross flows is search frictions. If workers are from time to time separated from their jobs and must search for new ones, and if for every worker there is some probability of finding a job in each sector, then there will be workers flowing in all directions, even in the steady state. This can occur either if search is undirected or directed, meaning respectively that workers cannot or can choose the sector in which they search.

Davisdson, Martin and Matusz (1999) incorporate search frictions of the Mortensen and Pissarides variety into a two-sector, two-factor trade model. Each country has a fixed amount of homogenous labor and capital, which combine to produce two goods. One unit of capital needs one unit of labor to produce, and each owner of unemployed capital must search for a worker, while each unemployed worker must search for a capitalist. Once joined, the worker and capitalist bargain over the surplus and then produce together until they are randomly separated and search begins again. Search is directed: Each searcher can choose
a sector in which to search. In equilibrium, unemployed factors search in both sectors, with
the proportions adjusting so that each worker and capitalist is indifferent between the two
sectors. This is the source of gross flows: A certain fraction of workers separated from
X employers happen to search in Y, and *vice versa*. The analysis shows that in steady
state, differences in the efficiency of search can be a source of comparative advantage, and
a generalization of the Stolper-Samuelson Theorem applies to searching factors. In general,
workers in the two sectors will receive different wages, but searchers in the two sectors must
receive the same utility. This genre of model is extensively explored in Davidson and Matusz
(2010).

This type of model would lend itself to dynamic modeling in principle, but almost all
of the available work focusses only on the steady state. The reason is that there does not
appear to be any analogue to the social planner’s problem in the bare-bones model discussed
above, since frictional-search models have generically inefficient equilibria. For this reason,
analytical results in this literature for the transition path – which are essential for the agenda
described at the outset of this review – are scarce.

One exception is Davidson and Matusz (2010, Chapter 2010), which studies a stylized
two-sector open economy with a low-tech, protected import-competing sector and a high-tech
export sector which requires both training and search for a worker to enter. Assumptions
are tailored to make search efficient, and then the transition path is solved in closed form.
The main finding is that once the parameters are calibrated to US data, the delay and costs
associated with moving to the high-tech sector take away about 40% of the gains from trade.
Another is Davidson and Matusz (2006), which studies a very simple and stylized two-sector
model in which entry to one sector requires search. They show that the present discounted
value of welfare is maximized by free trade, and the steady state of that equilibrium does
not maximize steady-state real income. This provides a crisp example of the importance of
the transition path in welfare analysis; focussing on the steady state can be very deceptive.
Importantly, in both of these models equilibrium search is efficient, which is unusual for the
search literature, and that is where the tractability comes from.

6The famed Hosios (1999) condition for efficiency is that at every date, in every state of nature, the
elasticity of the number of matches with respect to the number of searching workers is the same as the
workers’ share of the bargaining surplus in negotiations with employers. In only a very special, knife-edge
case will this hold.
Itskhoki and Helpman (2014) study a model with two sectors, a Ricardian sector producing a homogenous good, and a monopolistically-competitive sector with heterogenous firms. Workers can search for a job in either sector, and bargain with an employer once they find a match. Following trade liberalization, there is reallocation with frictional unemployment between sectors but also within the differentiated sector, as less-productive firms shrink or exit and more-productive firms expand.

Coşar (2013) studies a rich search-based model of labor adjustment, calibrated to data from Brazil. There are two sectors; sector 1 exports, while sector 2 is import-competing and is protected by a tariff. Production requires a match between a worker and a firm, following frictional search. The output generated is a function of (i) the sectoral production function; (ii) the sector-specific human capital possessed by the worker; and (iii) the quality of the match, which is random but remains constant as long as that worker remains with that firm. Over time, as long as a worker works in sector $i$, her sector-$i$ human capital accumulates. This is transferable across firms in the same sector, but is not useful in the other sector. Worker-firm relationships break up for exogenous reasons, but also for a very subtle endogenous reason: Human capital and the quality of a match combine multiplicatively, so as a worker gains experience in a sector, over time she has more to gain from upgrading the quality of her employment match, and may choose to leave a mediocre match in search of a good one.

Search is undirected: Workers cannot choose one sector in which to search. For that reason, workers recently departed from a sector-1 firm may accept a job offer from a sector-2 firm, and vice versa. This is the source of gross flows in excess of net flows in this model.

The model is calibrated to the Brazilian labor market in order to study the trade liberalization of 1991. A major finding is that adjustment to the new steady state following the liberalization is not merely slow, but inefficiently slow: A temporary subsidy to moving from the previously protected import-competing sector to the export sector speeds up the transition and increases the present discounted value of welfare. The reason is that the search structure creates a positive externality from switching sectors. Moving to the export sector amounts to an investment decision. The worker abandons her accumulated human capital in one sector to begin to acquire human capital in the other sector. Because at some point in the future she may drop her first export-sector employer and move to a new one, a portion of the benefit from this investment will accrue to a future employer, whose identity is
unknown at the time the worker contemplates switching sectors. As a result, fewer workers will make the switch than would be socially optimal.

This finding underlies a crucial point: How one thinks about the welfare effects of the adjustment process depends on how one models it. The bare-bones model discussed above and its cousins in the literature are neoclassical; the adjustment process does not add any inefficiency. The social-welfare maximizing policy, maximizing the sum of the net present value of expected utility across all citizens of a small open economy, would be a sudden move to free trade, with no government intervention in the resulting gradual adjustment in the labor market. However, frictional search confers market misallocation that in general implies that either a tax or a subsidy to adjustment will be optimal, and failing that, free trade will generally be suboptimal even in a small open economy. The two ways of modelling gradual adjustment with gross flows in excess of net flows are not simply two ways of achieving the same end, but are fundamentally different ways of thinking about the adjustment process.

Firing costs. Another approach to modeling dynamic labor adjustment to globalization focusses on employment regulations that slow down reallocation of workers. Kambourov (2009) studies the effect of firing costs, which are payments that must be made by an employer when dismissing a worker, on labor reallocation after a trade liberalization. This is a common feature of labor law particularly in Latin America, and is justified by its supporters as a way of protecting workers from capriciousness by employers and from labor-market risk (the policy is often called ‘employment protection’). In this model, a large number of perfectly-competitive industries produce output using sector-specific human-capital that workers accumulate with experience. The human capital is not transferable across industries, and workers are paid their marginal value product, so a worker receives a higher wage ceteris paribus as she gains experience in an industry. Each industry undergoes technology shocks over time, and if an industry’s technology shock is bad enough, less-experienced workers will leave it and search for a job in another industry, starting in the new industry at the bottom of the experience ladder. Search is undirected, and this is the source of gross flows.

Firing costs are a tax that an employer must pay whenever a worker is separated from her job. Kambourov points out that for equilibrium values it does not matter whether the employer or the worker must pay the tax, so he models it as a tax that the worker must pay; one could also call it a ‘quitting cost.’ The tax is higher for a more experienced worker,
and there is no firing cost for a new worker. Thus, there are three industry-switching costs a worker must consider before leaving her current industry: The direct firing cost; the one period of lost wages due to time spent searching; and the loss of industry-specific human capital. The model is calibrated to Chilean and Mexican data to study the importance of firing costs for the effects of trade liberalization. In both countries, firing costs are found to slow down reallocation and lower the welfare gains from the reform.

Coşar, Guner and Tybout (2016) put all of these elements together, studying intra-sectoral reallocation with both firing costs and frictional search, in a model with heterogeneous firms with persistent productivity shocks. The structural parameters are estimated on Columbian data using the method of simulated moments. Trade liberalization has multiple effects on worker-level labor dynamics. Lowering tariffs allows the more productive firms to import inputs, grow and expand into export markets, while less-productive firms shrink or exit. Employment therefore becomes more concentrated into large firms, which are unlikely to exit when they receive a productivity shock, but are more likely to adjust hiring in response to a shock. The net effect is a small increase in turnover. In addition, wage inequality rises, as disparity in rents across firms rises; and unemployment rises, as the desirability of searching in the manufacturing sector goes up.

A final device for modeling labor dynamics is generational change, as old workers exit and new ones enter, choosing their first industry optimally. This is a feature of the Dix-Carneiro (2014) and Traiberman (2016) models below, for example, but Matsuyama (1992) isolates that one mechanism in an elegant way. In that model, workers cannot reallocate once they have chosen a sector, so the dynamic adjustment to a trade shock comes entirely through new labor market entrants.

3 Structural empirical approaches.

Even the simplest models suggest that answering the primal question identified at the outset will be difficult, and likely impossible in an empirical study that observes how wages respond to changes in trade policy. In the bare-bones model, for example, it is easy to construct examples in which the real wage in the import-competing industry falls when the tariff is removed and creeps upward but never recovers its original value – and yet import-competing
workers benefit from the reform. The reason is option value (recall (4)). The real wage rises in the other sector, and each import-competing worker knows that each period there is a chance that she will move and take advantage of it. In this case, merely showing the short-run and long-run response of wages to the liberalization will not help in identifying welfare effects. For another example, it could be that the import-competing real wage falls in the short run but rises to a higher value than its original value in the long run, and yet the worker’s lifetime utility falls because the transition path to the higher wages takes too long (and the option value along the way is insufficient compensation). In this case, the long-run wage response is inadequate to answer the question of welfare. For these reasons, the literature has focussed on structural models, in which welfare can be analyzed through simulations.

In principle, one could estimate the structural parameters on panel data of workers by solving the equilibrium time-path of each worker’s career for each guess for the parameter vector, and then choosing the parameter vector to maximize the likelihood. In the labor literature, Keane and Wolpin (1997) use simulated method of moments to estimate their model of occupational choices, for example, and Lee and Wolpin (2006) use simulated method of moments in an analogous manner. In the trade literature, however, several alternative techniques have been employed, with complementary advantages. We turn to those now.

3.1 First wave: No unobserved heterogeneity.

(i) A first attempt. Artuç, Chaudhuri and McLaren (2010) (henceforth, ACM) took a slightly enriched version of the ‘bare-bones’ model to US data, with $n$ sectors and a state vector $S$ that evolves over time according to a Markov process. This random aggregate state can represent foreign export shocks, domestic shifts in labor demand due to technology shocks, and changes in policy – or perhaps signals of future changes in policy. The idiosyncratic preference shocks $\varepsilon^i_t$ that create gross flows are assumed to be distributed as Type I extreme-value with parameters constrained so that the mean shock is zero. The parameter $\nu$ is proportional to the standard deviation of the $\varepsilon^i_t$ shocks. In the main specification, the common cost of switching sectors is assumed to be the same for any origin-destination pair, and is denoted $C$.

The estimation method centers on an equilibrium condition analogous to the Euler equa-
tion in optimal savings problems. Using notation analogous to Section 2.2, we use $u^i(L_t, s_t, \varepsilon_t)$ to denote the worker’s maximized utility conditional on the current allocation vector $L_t \in \mathbb{R}^n$, the aggregate state $s_t$, and the vector of $n$ idiosyncratic shocks $\varepsilon_t$; $v^j(L_{t+1}, s_{t+1})$ the expectation with respect to $\varepsilon_{t+1}$; and $C^{i,j}$ the common value of the cost of switching from sector $i$ to $j$. Then the worker’s Bellman optimization condition can be written as:

$$u^i(L_t, s_t, \varepsilon_t) = w^i_t + \max_j \{\varepsilon^j_t - C^{i,j} + \beta E_t[v^j(L_{t+1}, s_{t+1})]\}$$ (5)

$$= w^i_t + \beta E_t[v^i(L_{t+1}, s_{t+1})] + \max_j \{\varepsilon^j_t + \bar{\varepsilon}^{i,j}_t\}$$

where:

$$\bar{\varepsilon}^{i,j}_t \equiv \beta E_t[v^j(L_{t+1}, s_{t+1}) - v^i(L_{t+1}, s_{t+1})] - C^{i,j}.$$

(6)

The term $\bar{\varepsilon}^{i,j}_t$ is analogous to the value $\bar{\mu}^i_t$ from Section 2.2, and can be interpreted as the ‘marginal idiosyncratic cost of moving’ from $i$ to $j$, or the idiosyncratic cost ($\varepsilon^j_t - \varepsilon^i_t$) of moving for a worker who is just indifferent between staying in $i$ and moving to $j$. As in the bare-bones model, the worker’s utility can be written as a current real wage, the value of continuing in the current sector, and an option-value term. Now, (5) can be taken to the data once it is transformed in two ways.

$$v^i(L_t, s_t) = w^i_t + \beta E_t[v^i(L_{t+1}, s_{t+1})] + \Omega(\bar{\varepsilon}^i_t),$$ (7)

where $\bar{\varepsilon}^i_t = (\bar{\varepsilon}^i_1, ..., \bar{\varepsilon}^i_N)$ and $\Omega(\bar{\varepsilon}^i_t)$ is the expectation of the option-value term with respect to $\varepsilon_t \equiv (\varepsilon^1_t, ..., \varepsilon^n_t)$. A simple transformation turns this condition into:\footnote{Substitute in (7) onto the right hand side of (6) to eliminate the $v^i(L_{t+1}, s_{t+1})$ term (but a $v^i(L_{t+2}, s_{t+2})$ is now added); take the difference between this equation for $i$ and $j$; and use (6) to eliminate the $v^j(L_{t+2}, s_{t+2}) - v^i(L_{t+2}, s_{t+2})$ term.}

$$C^{i,j} + \bar{\varepsilon}^{i,j}_t = \beta E_t[w^i_{t+1} - w^i_{t+1} + C^{i,j} + \bar{\varepsilon}^{i,j}_{t+1} + \Omega(\bar{\varepsilon}^j_{t+1}) - \Omega(\bar{\varepsilon}^i_{t+1})].$$ (8)

This is an Euler equation: It equates the cost of moving one more worker from $i$ to $j$ at time $t$ with the marginal benefit of doing so. On the left hand side, $C^{i,j}$ is the common portion of the cost of moving one more worker, and as discussed above, $\bar{\varepsilon}^{i,j}_t$ is the idiosyncratic portion.
The right hand side has the marginal benefit in three parts: The difference in next-period wages; the difference in next-period continuation values (proxied by $C^{ij} + \bar{\epsilon}_{ij}^{t+1}$ due to next period’s Euler equation), and the difference in option values.

Not only is (8) an equilibrium condition, but it is also an estimating equation. From the functional forms, we can derive that:

$$\bar{\epsilon}^{ij}_{t} = \nu [\ln m^{ij}_{t} - \ln m^{ii}_{t}]$$

and

$$\Omega(\bar{\epsilon}^{i}_{t}) = -\nu \ln m^{ii}_{t}. \tag{9}$$

As a result, all of the terms in (8) are either potentially observable (such as the wages and the gross flows $m^{ij}_{t}$) or a parameter to be estimated (namely, the $C^{ij}$ terms and the volatility of idiosyncratic shocks, $\nu$). Further, the estimation is extremely easy, since it becomes essentially a linear regression equation:

$$\ln m^{ij}_{t} - \ln m^{ii}_{t} = -\frac{(1 - \beta)}{\nu} C^{ij} + \frac{\beta}{\nu} (w^{j}_{t+1} - w^{i}_{t+1}) + \beta (\ln m^{ij}_{t+1} - \ln m^{jj}_{t+1}) + \xi^{ij}_{t+1}, \tag{10}$$

where $\xi^{ij}_{t+1}$ is a forecast error revealed at date-$t + 1$ which has a mean zero conditional on all public information available at date $t$. This forecast error will generally be correlated with date-$t + 1$ variables, so instrumental variables are required, and past values of endogenous variables can be employed.

The fact that the parameters of interest can be recovered simply, without any recursive computation of the equilibrium, may be surprising. Part of the reason is that the recursive nature of the equilibrium is used to eliminate the next-period-ahead value functions. For this, it is crucial that the workers have no unobservable heterogeneity; all workers must perceive the future attractiveness of all sectors in the same way. Beyond that, it is worth underlining why (10) identifies the desired parameters. For the moment, ignore $\beta$. Roughly, (10) relates today’s gross labor flows (the left-hand side) to tomorrow’s expected wage differential and tomorrow’s gross flows, with an intercept. If there are large amounts of gross flows at all times, a high value of the intercept is indicated (close to zero, that is), which implies a low value of $C^{ij}$ or a high value of $\nu$. Either the average cost of moving is low, or workers’ passions for other things in life are great, so that they chase those passions around and switch sectors often. At the same time, if today’s gross labor flows respond strongly to a change in tomorrow’s expected wage differential, then a large slope coefficient and hence a low value
of $\nu$ is indicated. In the limit, if $\nu$ becomes very small so that people care almost entirely about wages, a small change in expected wage differential will induce a huge movement of people. Thus, both the overall level of gross flows and their sensitivity to wage differentials suffice to identify these two structural parameters.

The model is taken to data from the US Current Population Surveys (CPS). The CPS is not a panel data set, but each March respondents are asked what their main job was in the previous year. Using these retrospective questions, a time series of the matrix of gross flows $m_{ij}^t$ can be constructed, and using the information on the respondents’ current job a time series of wages can be constructed. The economy is aggregated into 6 sectors, and the time span is from 1975 to 2000. As a result, equation (8) is estimated with $6 \times 5 = 30$ gross flows, and 24 years of useful data (accounting for lags for instruments), for a sample size of 720. The results indicate very high switching costs; in the preferred specifications (shown in panel IV of Table 3 and in Table 5), the common portion $C_{ij}$ of the cost amounts to between 4 and 7 times average annual income for most transitions. At the same time, the volatility of idiosyncratic shocks is also estimated to be large, in the neighborhood of 5 times average annual wages, as is needed to generate gross flows at the level seen in the data.

The model is used to simulate a trade liberalization of the sort described in the bare-bones model above, with a high tariff on manufactures suddenly removed. The economy takes about 8 years to approach its new steady state closely. The liberalization leads to an abrupt drop in the real wage in manufactures, but an abrupt rise in the real wages in all other sectors due to the drop in the domestic price of manufactures. A movement of workers out of manufacturing begins, and slows until the new steady state, with the real wage creeping upward in manufacturing and downward everywhere else as the supply of labor moves in each sector. The real manufacturing wage never recovers its pre-liberalization value. Crucially, despite this, the welfare of a worker employed in manufacturing at the date of the surprise liberalization rises. The reason is option value: The rise in the real wage in all other sectors creates enough of an attractive promise that the loss of manufacturing wages is outweighed.

(ii) Refinements: Using Conditional Choice Probabilities. Many weaknesses of ACM were addressed in Artuç and McLaren (2015). First, equation (8) cannot be used if there are zeros in the flow matrix (that is, a value of $m_{ij}^t = 0$). This forces a high degree of aggregation (at least with a modest-sized dataset such as the CPS), and in ACM the state
space is squeezed into 6 sectors. However, the problem can be cast in a form that makes use of predicted gross flows instead of the log of gross flows, circumventing this problem. In particular, the Poisson Pseudo Maximum Likelihood (PPML) approach suggested by Santos Silva and Tenreyro (2006) for estimation of gravity models lends itself well to this approach. Relaxing the zero-flows constraint allows for a much richer disaggregation of the data, and allows, in this case, for both sectoral switching and occupational switching. Using the same data, the economy is broken into 4 sectors and 5 occupations, for a total of 20 ‘cells,’ and this is done for college-educated and non-college-educated workers.

Second, implementation of the PPML approach is made possible by an scheme very closely analogous to the Conditional Choice Probability (CCP) method of Hotz and Miller (1993), which has become widely used in the micro-econometrics of dynamic choice. The core insight of CCP methods is that in a dynamic choice problem, the optimal choice taken at date \( t \) depends on the payoffs that the individual expects to realize at date \( t + 1 \) from the different actions that might be taken at \( t \). In equation (5), these are the expected \( v^j \) terms. Of course, these are unobservable to the econometrician (and will tend to be correlated with observables; for example, an industry just liberalized will tend to have a low \( w^j \) and a low next-period \( v^j \)). However, an option that is generally very attractive will tend to be chosen by a lot of people, so one can use observed probabilities that a person chooses a given option to infer information about the future expected payoffs. In these models, the probability that a worker of a given type takes an action is observable (with enough observations) as the fraction of workers of that type who take that action – the gross flows.

In this case, the estimation method takes two steps. First, given the extreme-value distribution assumption, gross flows between any two sector-occupation cells at any date can be written:

\[
m_t^{ikjls} = \frac{\exp \left[ \frac{1}{\nu} \left( \beta E_t \left( v_{t+1}^{jls} - v_{t+1}^{iks} \right) - C(i, k, j, l, s) \right) \right]}{\sum_{j'=1 \ldots I, l'=1 \ldots K} \exp \left[ \frac{1}{\nu} \left( \beta E_t \left( v_{t+1}^{j'l's} - v_{t+1}^{iks} \right) - C(i, k, j', l', s) \right) \right]},
\]

where \( m_t^{ikjls} \) denotes the fraction of workers of type \( s \) in sector-occupation cell \( (i, k) \) (that is, sector \( i \) and occupation \( k \)) who choose to move to cell \( (j, l) \) in period \( t \); \( v_{t+1}^{jls} \) is the payoff for a worker in cell \( (j, l) \); and \( C(i, k, j, l, s) \) is the common cost of switching from cell \( (i, k) \) to \( (j, l) \) for a worker of type \( s \).
The denominator of (11) is a factor that is the same for all destination cells \((j', l')\), conditional on starting from the cell \((i, k)\). The terms in the numerator are: one that depends on the destination \((j, l)\); one that depends on the origin \((i, k)\); and one, \(\frac{1}{\nu} C(i, k, j, l, s)\), that depends on both. This structure allows us to choose values of \(\frac{1}{\nu} \beta E_{t} v_{t+1}^{iks}\) for all \((i, k)\) and \(s\) at each date, as well as parameters of switching-cost function, to match the empirical rates of gross flow as closely as possible, following the PPML framework. This is the application of the CCP idea, that observed choice probabilities can provide a great deal of information about expected future payoffs from the various options.

The second step is to use the worker’s Bellman equation, with \(\frac{1}{\nu} \beta E_{t} v_{t+1}^{iks}\) and \(\frac{1}{\nu} C(i, k, j, l, s)\) treated as data, to estimate \(\nu\). The core idea is, as in ACM, that if gross flows (and therefore \(\frac{1}{\nu} \beta E_{t} v_{t+1}^{iks}\)) are highly responsive to changes in next-period expected wage differentials, a low value of \(\nu\) is indicated.

An additional refinement relative to ACM is that each cell is allowed to confer a non-pecuniary benefit that is common to all workers, which can allow for a cell to remain popular with workers even if its wage is low. This had already been shown to be important in Dix-Carneiro (2014) (discussed below), and tends to reduce estimated switching costs, because in their absence workers remaining in a low-wage sector would be attributed to those costs.

The estimated moving costs are again quite large, generally in the range of 3 to 5 times annual income for the common portion. They are generally similar in magnitude for a switch from one occupation to another within a sector; from one sector to another within an industry; or in a switch in both dimensions at once. The standard deviation of the idiosyncratic shock is about 4/5 of annual income.

The paper simulates two separate shocks (plus a combination of the two), a trade shock as in ACM, and also an offshoring shock, meaning a drop in the cost of using foreign production workers for domestic production. The simulation results are much more pessimistic than they were under ACM. (i) With the estimated parameters, the trade shock lowers the expected lifetime utility of all non-college-educated workers who are initially employed in manufacturing, regardless of their occupation. Option values are important, but with this refined structure it turns out that they are not enough to improve outcomes for that most affected group. However, every other group of workers in both educational classes benefits. Note that college-educated workers in manufacturing suffer a dramatic drop in real wage, in
all occupations, at the date of the liberalization, but their long-run real wage is slightly higher than the pre-liberalization wage, and their improved option value is enough to make them barely better off. Once again, one needs a dynamic model to evaluate the welfare effect. (ii) The offshoring shock induces manufacturers to shift production-worker labor demand toward foreign workers and away from domestic workers, pushing down wages for manufacturing-sector production workers for both educational classes. This induces a movement away from the manufacturing-production-worker cell, but there are far more non-college-educated than college-educated workers in the cell to begin with, so this migration pushes non-college-educated worker wages in all other sectors down over time, while this lowering of blue-collar wages throughout the economy raises college-educated workers’ wages over time. In the end, most blue-collar workers’ welfare falls, while almost all college-educated workers’ welfares rise. Ebenstein et al (2014) found in reduced-form regressions that blue-collar workers in occupations with routine-task-intensive occupations is US manufacturing tended to move into service sectors when their industries increased offshoring to low-wage economies. This result is consistent with that, but here we interpret it in welfare terms, and also find an effect that they cannot: The indirect downward pressure on all blue-collar wages through the dynamic response to the offshoring shock.

It is useful to return to our discussion of Caliendo, Dvorkin and Parro (2015) from Section 2.3. Recall that they found that, in every sector-state cell, the expected lifetime utility of a worker goes up at the date of the revelation of the China shock, relative to the pre-shock steady state, even for workers initially in the most vulnerable industries or states. However, this result may not hold up, just as the original finding in ACM did not hold up, once a distinction is made between high-skilled and low-skilled workers. It could be, as in Artuç and McLaren (2015), that once that distinction is permitted in the model, low-skilled workers in the most-affected sectors or locations will be worse off. Nonetheless, the contribution of Caliendo et al (2015) is substantial because the machinery in the paper makes such an inquiry possible.

Finally, Brussevich (2016) applies the PPML-CCP method to study the gender effect of the growth of Chinese exports on US workers. She estimates a switching cost function for occupations, based on differences in task characteristics across occupations as recorded in the O*NET occupational data, that is allowed to vary by gender. The switching-cost functions
turn out to be quite different by gender, and the finding that female workers are better able to switch into service occupations than male workers helps explain recent narrowing of the wage gender gap as a response to the China shock.

(iii) A shortcut for developing-country data. The above approaches require a time series of the full matrix of gross flows between all sectors, occupations, regions, or cells of interest to the researcher, as well as a time series of wages in each of the same. This type of data is not available for many developing countries, so the usefulness of this approach for policy analysis outside of the OECD is limited. However, Artuç, Lederman and Porto (2015) have concocted a kind of shortcut that allows estimation of the moving-cost parameters of this class of model with much less data.

The UNIDO dataset from the United Nations provides unified labor-market data for manufacturing across a wide range of countries, and Artuç et al extract from it a time series for each country of both the vector of employment levels by industry and the vector of wages by industry. (Gross flows across industries are not available.) They take as given the allocation of labor in each country in the first year of data, and also the time path of real wages in each industry for the time span of the data. They also impose an assumption that the value of $\nu$ is the same for all countries, and use a value that they estimate for the US from the CPS. In each country, free parameters are the common moving cost $C$ and non-pecuniary fixed effects for each industry $\eta_i$. Then, treating the horizon of each worker as finite (given by the time span of the data), for any guess of the parameter values they simulate the time path of the allocation of workers in a given country by recursively applying (4) and (9) with the appropriate analogue of (11). This time-path for the allocation of labor can be compared with the actual time path in the data, and a value of $C$ can be chosen for that country to minimize a weighted sum of square deviations between the simulated and actual allocation. This provides a consistent estimator for $C$ and the $\eta_i$ values for each country, given the assumptions of the model (assuming, of course, that the imposed value of $\nu$ is right).

This method yields estimates of $C$ for 56 countries, in each case scaled to a multiple of average annual real income. The estimates range from 1.29 (for Estonia) to 5.07 (for Jordan). There is a strong negative correlation between income per capita and $C$. The moving costs tend to be lower in countries with higher-quality labor-market institutions, better contract
enforcement, and better trade facilitation. These correlations suggest that some portion of $C$ may be a result of the institutional and policy environment, and so could be lowered by improvements in that environment. The paper simulates a trade shock that takes the form of an abrupt drop in the price of output in the food and beverage sector, which could result from elimination of a tariff or a foreign supply shock. In each case, real wages in the food and beverages sector drop on impact, then creep upward as workers leave the sector, arriving at a steady state with a higher real wage in the sector than before the shock. In almost every country, the welfare of workers initially in the food and beverage sector rises as a result of the shock. The length of the transition path varies greatly, from 2 years for low-$C$ countries to 12 years for high-$C$ countries.

(iv.) Some benefits of these approaches. The approaches described in this section are very different, but have two very attractive features in common (aside from their computational economy). (1) They have very modest data requirements. The first two do not technically require individual data on workers at all, but only aggregate rates of gross flow and average wages. Now, normally these are derived from individual data, but even so, they do not require longitudinal data that follows individual workers’ careers over time. The last approach obviously has trivial data requirements. (2) These approaches confer what we might call an agnostic dividend. Since they do not require future values of the value function to be computed, they do not require that the researcher make any assumption about what the workers know or believe about the future, only that the expectations about date $t + 1$ events are unbiased, conditional on all information available publicly as of date $t$. If, in the time span of the data, newspapers report that the governments of the US and Mexico are making good progress toward a free trade agreement, or that the president is making good progress persuading Congress to grant Permanent Normal Trade Relations with China, that can radically change expectations about the time path of wages in affected industries or occupations. That causes no problem with these estimation methods. In addition, the researcher need know nothing at all about the production side of the economy and labor demand in order to estimate the parameters of workers’ mobility costs. This is of course not the case with policy simulations – at the least strong assumptions about the production side must be imposed for those.
3.2 Approaches with unobserved heterogeneity.

As noted above, assuming away unobserved heterogeneity is attractive because it requires little computing power, is transparent and easy to understand, and can be applied to very limited data sets, making it particularly useful for developing-country data. It also yields very tractable theory. However, the drawbacks are extremely obvious. The assumption is implausible, and it could lead to an overestimate of the moving-cost parameters. If a worker has a comparative advantage in sector $i$, she could choose to remain in that sector even if its wages fall, since she knows that her own earnings in any other sector would be much lower. The limited responsiveness of labor flows to wage differentials observed in the data could be incorrectly attributed to high switching costs when it is in fact just a result of unobserved worker characteristics.\footnote{This might not change the conclusions on the distributional effects of trade, however. If some workers stay where they are even after their incomes have been depressed by a trade shock, the exact source of their immobility does not matter for the conclusion that their lifetime earnings have been reduced.} A rich strain of work has been developing that abandons this assumption completely to embrace much richer and more realistic treatment of mobility costs.

Dix-Carneiro (2014) specifies a rich model of labor mobility in a trade model and estimates it on Brazilian data. Its description of workers is orders of magnitude more realistic than the papers in Section 3.1. To begin, workers are finite-lived, with a working life from age 25 to 60. Each sector $s$ confers a common non-pecuniary benefit to its workers, $\tau^s$, allowing for a permanent compensating wage differential. Each worker $i$ also realizes a time-varying, iid idiosyncratic preference shock $\eta_{it}^s$ for sector $s$, providing for gross flows as in ACM. The crucial differences are in the specification of switching costs and human capital. The cost of switching from one sector to another is specified as a function of the worker’s gender, education, and age (in quadratic form), as well as an unobservable worker fixed effect that allows for the possibility that some workers are just better able to adjust and move than others.

Production in each of the 7 sectors is a function of high-skilled labor, low-skilled labor, and capital. In each case it is not raw labor that matters, but total human capital. Each worker’s wage in a sector is equal to her sector-specific human capital times the price of that type of human capital in that sector, which is equal to the marginal value product of that type of human capital in that sector. Human capital accumulates endogenously with work.
experience. Human capital for a low-skill worker \( i \) in sector \( s \) is specified as follows:

\[
\begin{align*}
    h^{0,s} &= \exp \left[ \beta_{s1}^i Female_i + \beta_{s2}^i I(Educ_i = 2) + \beta_{s4}^i (a_{it} - 25) + \beta_{s5}^2 (a_{it} - 25)^2 \right] \\
    &\quad + \sum_{k=1}^{7} \beta_{s5+k}^i Exper_{ikt} + \theta_i^s + \epsilon_{it}^s 
\end{align*}
\]

Here, the superscript 0 on the left-hand side indicates that this is the equation for type-0 workers, or unskilled workers, defined as less than high-school graduates (the equation for high-skilled workers is analogous). The variable \( Female_i \) denotes gender; \( I(Educ_i = 2) \) is a dummy variable for some middle school but no high-school graduation; \( a_{it} \) denotes the worker’s age, so \( (a_{it} - 25) \) is the number of working-age years the worker has had so far; \( Exper_{ikt} \) is the number of years the worker has had so far in sector \( k \); \( \theta_i^s \) is the effect of the worker’s unobservable type, indicating the degree of talent for sector \( s \); the \( \beta_k^s \) are coefficients to be estimated; and \( \epsilon_{it}^s \) is a random disturbance. There are two crucial features to note. First, the \( \theta_i^s \) term is observable to the worker but not to the econometrician, and so obliterates the possibility of using the Euler-equation type methods discussed above. Second, the specification allows for partial transferability of work experience across sectors. If \( \beta_5^s = 0 \), for example, then experience in sector 5 is useless in sector \( s \), but if \( \beta_5^s = \beta_s^s \), then experience in sector 5 is just as useful in sector \( s \) as experience in sector \( s \) would be. This matrix of experience terms is estimated along with the other parameters, and is an important component of moving costs: The human capital that a worker loses when switching sectors can be an important inducement to stay put.

As a result, this model has three types of mobility cost: The explicit costs (common as well as idiosyncratic, as in the bare-bones model), allowed to be a function of worker characteristics but still essentially a black box; the loss of human capital from switching sectors; and the worker-specific comparative advantage terms \( \theta_i^s \). The estimation procedure allows us to see the relative importance of these elements.

The model is brought to the RAIS panel data, which is a matched employer-worker data set that includes all formal-sector employed workers in Brazil. The data used here span 1986 to 2005. The model is estimated by indirect inference. For each candidate value for the parameters, the equilibrium is solved and used to generate a simulated data set. The
researcher chooses a number of regressions to run on both the real data and the simulated data, whose coefficients can be interpreted as important moments of the data, and parameters are chosen to minimize the distance between the coefficients estimated on the real and simulated data sets. Examples of these regressions used in this case are regressions for which the dependent variable is the sectoral wage or residual wage variance, and linear-probability regressions for a worker’s sectoral choice based on current conditions as regressors.

The results reveal that all three elements of switching costs are important. The unobserved types are limited to three values, and the estimates show that the difference in abilities (and in moving costs) across the types are large. Non-transferability of human capital is a large impediment to moving; in most sectoral transitions, accumulated experience is worth less than half of what the same experience would have been worth if it had been accumulated in the same sector. In addition, direct, common costs of switching sectors are large, but much smaller than in ACM: Most switches from one productive actor to another cost between one and two year’s wages.

The trade-liberalization simulation is a sudden elimination of a 30% tariff protecting the high-tech sector, which is an import-competing sector in Brazil and still enjoys substantial trade protection. The simulation is performed under a range of assumptions regarding capital mobility (a significant innovation), and under all cases the adjustment is sluggish, with the economy taking at least 9 years to reach 95% of the distance to the new steady state, and welfare for workers in the high-tech sector falls substantially. Under perfect capital mobility, the welfare of high-tech sector workers falls by 9.9% on average. Very importantly, the loss is much larger for older workers and for high-skilled workers than for young or high-skilled. For old skilled workers the loss is 22.5%, while for young and skilled it is 6.5%. This is both due to the shorter time horizon for older workers, making it less likely that switching sectors will be optimal, and to higher moving costs estimated for older workers.

Subsequent to the Dix-Carneiro Brazil study, Ashournia (2015) used the framework with a few modifications to estimate mobility costs in Denmark. The estimated common costs of switching sectors was dramatically less, ranging between 10% to 20% of annual income. However, the cost of lost human capital in switching sectors was found to be large, and the adjustment to a trade shock still requires more than 10 years to get close to the new steady state.
The advantages of this approach are obvious. Some disadvantages are more subtle. First, this kind of richness requires a large and detailed panel data set on individual workers. The RAIS data is not only unusually suited for this purpose, but is somewhat extraordinary. More subtly, because the estimation method requires actually simulating an equilibrium (many times), it is necessary to specify what information and expectations workers have as part of the estimation. So we lose the ‘agnostic dividend’ discussed above (Section 3.1 (iv)). Further, in order to speed the process of finding an equilibrium, an ad hoc expectations function is specified for workers, as is done in Lee and Wolpin (2006). Workers are assumed to believe that the first differences of the human-capital prices in each sector follow an AR(1) process, whose parameters are estimated together with the rest of the model’s parameters. The coefficients of this forecast rule are consistent, in the sense that the behavior they induce in workers generates a time-series process for wages that is consistent with the forecast rule. However, it is at best an approximation to a rational-expectations equilibrium. Note, importantly, that if in the course of the data there is any change in the political environment that makes a future change in trade policy more likely – an election, a proposal in the Congress, the Doha round of the WTO process which got underway in the middle of the data set – it would not be captured by this forecasting rule.

Traiberman (2016) pursues an alternative approach to the unobservable-heterogeneity problem. His focus is on switching occupations rather than industries, but the same issues arise. Parallel with Dix-Carneiro (2014), a worker’s wage is determined by her occupation-specific human capital times the human-capital price in that occupation at the moment, and a worker’s human capital is a function of age, experience, skill class, an unobservable type or ‘talent shock,’ and an idiosyncratic, time-varying productivity shock. The unobservable ‘talent shock’ is time-invariant and denoted $\theta_{oi}$ for worker $i$ in occupation $o$. In a major contrast with Dix-Carneiro, experience is assumed to be completely untransferable across occupations, as in Coşar (2013) and Kambourov (2009), in order to focus on quantifying occupational switching costs. Those costs are specified as a function $f(\omega)C(o, o')$ of the initial occupation $o$, the destination occupation $o'$, and worker characteristics $\omega$ both observed (age and education) and unobserved ($\theta_{fi}$). The function $C$ is a function of the task composition of the occupations in question, based on the O*NET occupations data.

To deal with the dependence of both occupational human capital and switching costs on
the vector $\theta_i$, Traibernman turns to a set of techniques adapted from Arcidiacono and Miller (2011), which adapts the Conditional Choice Probability (CCM) approach explored by Hotz and Miller (1993) to the case of unobserved heterogeneity.

The technicalities are rather elaborate, but two main ideas are adapted for Traiberman’s empirical strategy. First, the EM algorithm (the acronym stands for ‘expectations-maximization’) can be used to manage the unobserved heterogeneity. Suppose that there are $K$ unobserved worker types, and the vector of unobserved characteristics $\theta$ possessed by each worker is the same for all workers of the same type. The fraction $q_k$ of the population who are of unobservable type $k$ and the value $\theta_k$ of the vector of unobserved characteristics that belong to workers of that type are both parameters to be estimated. The researcher does not know which worker in the data belongs to which type, but suppose that for each worker there is a Bayesian subjective probability attached to each type. Then, if there is a current guess as to the value of all parameters in the model, Bayesian updating can be used to update the probability assessed for each type for each worker, based on the action that the worker takes at each date. (Other things equal, a worker of a type that is specially talented at occupation $o$ will be more likely to choose that occupation, so observing her choosing $o$ raises the inferred probability that she is of that type.) Then the probabilities for type $k$ can be averaged across workers to provide an updated guess for the population share $q_k$. Now, with an updated probability for each type for each worker, it is straightforward to write the likelihood (following the second main idea, discussed below), and maximize it with respect to the parameters, holding the worker-specific probabilities for the unobserved types as well as the $q_k$ values as given. Iterations on this EM procedure – updating the Bayesian expectations, the maximizing the likelihood – converge to a consistent estimator of the parameters.

The second main idea is to use the structure of the model to simplify the construction of the likelihood. If one can write the probability of a worker choosing a given sequence of occupations at a given date as a function of all of the parameters, that can be used to construct a likelihood function, but because of the dynamic nature of the model that is generally not possible. Consider the probability, $\pi(\omega, o, o')$, that a worker in occupation $o$ at time $t$ chooses occupation $o'$ in the subsequent period. Due to the assumed generalized extreme-value distribution for the idiosyncratic shocks, conditional on the worker’s state
that probability can be written in a form similar to equation (11) above. It depends on parameters to be estimated but also on the expected payoff for that worker from being in each of the possible occupations at the start of the next period. These payoffs are of course not in the data, and because of the unobserved heterogeneity of workers they cannot be eliminated using the methods of Section 3.1. The trick used here follows from the fact that switching occupations in this model is what Arcidiacono and Miller (2011, Section 3.2.1) call a renewal action, an action such that, if two workers of initially the same state both take this action after any history, the distributions of their future states coincide. In other words, a renewal action wipes out differences in the history of actions between two otherwise identical individuals, once both individuals take that action. In this model, switching occupations is a renewal action because of the assumption that occupational experience is not transferable. Anyone of a given type \( k \) in his first day as a receptionist is a neophyte, equally unskilled in the task regardless of what else he had done in previous jobs.

To illustrate how this helps, consider two workers, both with initial state \( \omega \), and both in occupation \( o \) as of date \( t \). Suppose that worker 1 switches to occupation \( o' \neq o \) at date \( t \); and then at date \( t + 1 \), when his updated state is \( \omega' \), he switches to occupation \( o'' \). Worker 2 stays with occupation \( o \) for a period, then in period \( t + 1 \) when her updated state is \( \omega'' \), she also switches to occupation \( o'' \). Both workers begin and end in the same place, and because both are new to \( o'' \) as of the beginning of date \( t + 2 \), they both have the same human capital at that date, and therefore the same state and the same future prospects. As a result, their future payoffs from all of the possible occupational choices they could be making are the same, and if we take the difference in logs between their probabilities of making these choices, the future-payoff terms cancel out. The result is an estimating equation as follows:

\[
\log \frac{\pi_t(\omega, o, o')}{\pi_t(\omega, o, o)} + \beta \log \frac{\pi_{t+1}(\omega', o', o'')}{\pi_{t+1}(\omega'', o, o'')} = \tilde{C}(\omega, \omega', \omega'', o, o', o'') + \frac{1}{\rho} \left( w_{o't} h_{o'\omega}(\omega) - w_{o't} h_{o\omega}(\omega) \right) + \frac{1}{\rho} (\eta_{o'} - \eta_o) + \xi_{o't}, \tag{15}
\]

where \( \tilde{C}(\omega, \omega', \omega'', o, o', o'') \) is the difference in switching costs over the two periods for the two paths (constructed from the \( f(\omega)C(o, o') \) function); \( w_{o't} \) is the price of human capital for occupation \( o \); \( h_{o\omega}(\omega) \) is the occupation-specific human capital in occupation \( o \) of a worker.
with state $\omega$; $\rho$ is a parameter governing the variance of idiosyncratic shocks; and $\xi_{ot}$ is a forecasting error.

Note that this is far more economical computationally than the indirect inference method, and allows for a much richer choice space. Dix-Carneiro collapses the Brazilian economy into 7 sectors; this paper allows workers to choose between 24 occupations. An additional benefit is that it restores the agnostic dividend by eliminating the need for computation of workers’ expected future payoffs. In addition, the model can be estimated for the workers’ parameters without making any assumption about the production side of the economy. However, it does need an exceptionally rich data set, because the method requires following individual workers with similar but not identical career paths. Note also that the method depends on assuming that occupational human capital is not transferable at all, while Dix-Carneiro showed that this is strongly rejected in his data.

Traiberman uses linked employee-employer data on Danish workers from 1997 to 2007 with about 2.5 million observations per year. Occupational switching costs are found to vary widely by worker characteristics and occupation, but the median is between 5 and 8 times annual income. Simulation of a trade shock that takes the form of the observed drop in import prices over the period (contrasted with simulation holding those prices fixed) shows a rapid move to the new steady state, but a drop in welfare for most workers, which varies greatly by occupation. The gains from the country’s improved terms of trade are entirely captured by capital.

3.3 Observations.

Rather than searching for the one right way of analyzing the effect of international shocks on labor market dynamics, it is likely more helpful to collect a toolkit, each tool being the best for some purpose. First, reduced-form static approaches (Autor, Dorn and Hanson (2013) and Kovak (2013), for example) are helpful in establishing prima facie evidence of barriers to switching, and reduced-form dynamic approaches (Autor, Dorn, Hanson and Song (2014) and Dix-Carneiro and Kovak (2015), for example) can establish useful stylized facts that the dynamic models need to address. But a complete welfare analysis requires a structural model. The CCP techniques applied under the assumption of no unobserved heterogeneity (Section 3.1) are useful when data is limited, and actually produce results surprisingly similar
to the more sophisticated methods, once cell fixed effects are allowed for. But unobserved heterogeneity is unquestionably demanded by the data, as Dix-Carneiro (2014) showed, so a full account requires treatment of that as well as experience effects and life-cycle effects. Here as well there are different tools to choose from – so far, methods that require extremely rich longitudinal data. The indirect inference method allows for maximal richness and minimal assumptions imposed, but is computationally demanding. The E-M approach with special assumptions that create a renewal action is more tractable, but the special assumptions are not likely to be literally true.

One point that emerges from all of these approaches is that in all of the estimations, we find consistently – in the US, in Brazil, and in Denmark – that industry and occupational switching appears to be very costly for workers, and as a result transition following a trade shock tends to be slow (at least 8 years to reach 95% of the way to the new steady state for most simulations). The transition path matters greatly for welfare. At the same time, gross flows and option value matter consistently as well.

4 Issues for further exploration.

In closing, it may be helpful to mention a number of questions that have been raised but not settled, or that need to be raised in the first place.

(i) A taxonomy of switching cost. Early papers treated the switching cost \( C \) as merely a black box, but it would be valuable to be able to identify what its main components are. Dix-Carneiro (2014) advanced the question substantially by showing that in the Brazilian data a large piece in practice is the non-transferability of human capital. Ashournia’s (2015) results seem to imply that in the Danish case most of the switching cost takes this form. At the same time, Artuç, Lederman and Porto (2015) show that moving costs across countries seem to be correlated with the quality of institutions; it would be of great interest, for example, to identify how much of the variation in switching costs comes from policies such as the firing costs that are the focus of Kambourov (2009).

Risk aversion. It is standard in this literature to assume risk neutral workers. This has two costs. First, obviously we cannot assess the risk effect on workers’ welfare due to globalization. But more subtly, we cannot say anything meaningful about savings behavior.
In the standard model, in order to get positive consumption in each period, we must assume that the real interest rate is always equal to the rate of time preference. Once that has been assumed, optimal savings behavior is indeterminate. Now, if risk aversion is introduced, the social-planner’s approach to analyzing equilibrium from Section 2.2 no longer applies, and the question of at least imperfect risk-sharing institutions comes up. But vast richness can potentially be added at the same time.

*Life-cycle effects.* The more sophisticated models have featured workers with finite lives, whose costs and productivity can vary with age and whose decisions can also vary over the life cycle because of their shortening time-horizon as they age. In Dix-Carneiro (2014), older workers in the import-competing sector lose more than younger ones as a fraction of their lifetime utility. However, these models assume that the payoffs end at retirement, which occurs exogenously at age 60, while the welfare effects of globalization on retirement consumption ought in principle to be part of the analysis. Artuç (2012) models retirement income as the flow income from a representative bundle of capital assets, and the value of this income rises as a result of trade liberalization. Consequently, there is a U-shaped relationship between age and the effect of globalization: Young workers mostly benefit because even if they are in the import-competing industry they have plenty of time to switch; old workers benefit because they mostly care about their retirement; but middle-aged workers in the import-competing industry are hurt. A more serious treatment of this, including endogenous life-time saving, could be enlightening – but it would require incorporating risk aversion, as noted above.

**References**


