Columbia University

Department of Economics
Discussion Paper Series

Mobility Costs and the Dynamics of Labor Market Adjustment to external Shocks: Theory

Stephen Cameron
Shubham Chaudhuri
John McLaren

Discussion Paper #:0102-53

Department of Economics
Columbia University
New York, NY 10027

April 2002

Mobility costs and the dynamics of labor market adjustment to external shocks: theory*

Stephen Cameron
Department of Economics
Columbia University

Shubham Chaudhuri
Department of Economics
Columbia University

John McLaren
Department of Economics
University of Virginia, Charlottesville

February 2002

Abstract

We construct a dynamic, stochastic rational expectations model of labor reallocation that is designed so that its key parameters can be estimated for trade policy analysis. A key feature is the presence of time-varying idiosyncratic moving costs faced by workers. As a consequence of these shocks: (i) gross flows exceed net flows (an important feature of empirical labor movements); (ii) the economy features gradual and anticipatory adjustment to aggregate shocks; (iii) wage differentials across locations or industries can persist in the steady state; and (iv) the normative implications of policy can be very different from a model without idiosyncratic shocks, even when the aggregate behavior of both models is similar. It is shown that the solution to a particular planner's problem yields a competitive equilibrium, thus facilitating the analysis and simulation of the model for policy purposes.

Keywords: mobility costs, gross flows
JEL categories: D33, F16, J60

*We are grateful to seminar participants at the Board of Governors of the Federal Reserve System, Koç University; Syracuse University, University College, London, the University of Virginia and the World Bank; to participants in the European Research Workshop in International Trade, July, 2000 and the Summer Institute of the National Bureau of Economic Research, August, 2001; and also to Bill Gentry, Ann Harrison, Glenn Hubbard, and Marc Melitz for comments. This research is supported by NSF grant 0080731.

1. Introduction.

The effect of a given change in trade policy is greatly affected by the costs workers may face in adjusting to it. This is especially true of the *distributional* effects of the change, but it extends also to the efficiency effects. For example, the effects of opening up a sector of the economy previously protected from import competition depend crucially on how easily the workers in that sector can find employment in other sectors. If geographic or sectoral mobility costs are high, the efficiency benefits are thereby reduced and the burden borne by those workers is increased. Analysis of the effect of trade on wages thus always requires the use of some assumption on the degree of labor mobility.[2] Further, the effects of immigration into a particular region of the country depend on how fluid labor is between that region and others, sand so the literature on labor-market effects of immigration has always required assumptions on the degree of mobility (see Borjas et. al. (1996), Slaughter and Scheve (1999)).[3]

The cost of labor reallocation is also a central issue driving the political economy of trade policy, as emphasized for example in static approaches by Magee (1989) and Irwin (1996), and in dynamic analyses of endogenous trade policy such as Staiger and Tabellini (1999) and McLaren (1997, 2002).

This paper proposes a workhorse model of equilibrium labor reallocation that is designed to address these policy questions head-on. It incorporates a number of features that are intended to make the model helpful in analyzing trade policy changes in particular, and to be consistent with the broad empirical features of the adjustment process. It also has the benefit that its parameters can be estimated econometrically, thus providing for more detailed policy analysis through simulation, a project which is being carried on in parallel with the theoretical exercise detailed here.

The model is an infinite-horizon dynamic stochastic model with rational expectations, in which from time to time random shocks may hit labor demand

[2] For example, specific-factors models and the Stolper-Samuelson approach have very different implications for the relationship between trade and wages, driven entirely by different assumptions about mobility costs; and the appropriate time horizon for measuring the labor-market effects of trade also depends on assumptions about mobility costs. See Slaughter (1998) for an extended discussion.

[3] For example, the differences between the Hecksher-Ohlin approach, the "factor-proportions analysis" approach, and the "area analysis" approach to the effects of immigration (Borjas et. al., 1996) are entirely driven by different assumptions about labor mobility. See Slaughter and Scheve (1999) for an extensive discussion.

either in a sector or in a region of the country (for example, changes in trade policy or terms-of-trade shocks). In response to these shocks, each worker at each moment may choose whether to remain where she is or to move to another sector or geographic location. If the worker moves, she will pay a cost that has two components: A portion that is the same for all workers making the same move, which is a parameter of the model and is publicly known; and a time-varying idiosyncratic portion. The latter is an extremely important feature of the model, because it generates all of the model's dynamics and allows for gross flows to exceed net flows. If individual situations can vary, one may find large numbers of workers moving in opposite directions at the same time, and this is indeed a prominent feature of the equilibrium of the model. This is important because empirically gross flows of workers across geographical locations and industries are substantially larger than net flows.

Many authors have proposed theoretical models of the dynamics of factor reallocation in response to a trade or policy shock (a number of the issues are reviewed in Neary (1985)). Mussa (1978, 1982) studies the dynamics of adjustment in a trade model, with capital as a quasi-fixed factor bearing convex adjustment costs. In both models, labor is either completely immobile (that is, labor faces infinite moving costs) or costlessly mobile (faces zero moving costs). The roles of capital and labor could easily be reversed to consider labor adjustment dynamics. Dixit (1993) studies a similar model with random trade shocks and a fixed cost to each reallocation, and Dixit and Rob (1994) consider fixed labor-adjustment costs in a model with random labor-demand shocks and risk-averse workers. Rappaport (2000) and Dehejia (1997) study models with strictly convex reallocation costs for labor. Matsuyama (1992) studies a model whose workers cannot reallocate once they have chosen a sector, so the dynamic adjustment to a trade shock comes entirely through new labor market entrants.

Finally, two important papers are particularly closely related to the model used here. Jovanovic and Moffit (1990) offer an approach based on a matching model, in which workers disappointed in the job-match with their employers search for a better match, and in each period some fraction of them move across sectors to do so. Topel (1986) studies the dynamics of geographic reallocation of labor using an equilibrium overlapping generations model with idiosyncratic moving costs.

Our theoretical model differs from all of the above approaches in two ways. First, we allows gross flows to exceed net flows, which is important given the empirical importance of gross flows highlighted above. Jovanovic and Moffit (1990) shares this feature, but the other studies mentioned above do not. For this reason, idiosyncratic shocks are a key feature of our model. Unlike Jovanovic and Moffit (1990), we allow for such shocks to be non-pecuniary in nature (such as job dissatisfaction or personal constraints on geographic location).[4] Topel (1986) allows for idiosyncratic moving costs, but constrains gross interregional flows to be equal to net flows.[5]

Second, our model has been tailor-made to allow for estimation of the moving-cost parameters, a feature shared by none of the other equilibrium models.

In examining the model, we first study a particular (distorted) planner's problem in some detail, because it turns out that the planner's solution is also a market equilibrium. This provides a number of results on the market equilibrium that would be very difficult to derive by other means. The key properties include gradual adjustment of the economy to an external shock; anticipatory adjustment of the economy to an anticipated shock; and persistent wage differentials (across sectors or regions of the economy) even in the long-run steady state, for reasons that appear to be novel in the literature. In addition, we discuss a thought experiment that demonstrates the importance of idiosyncratic costs for empirical work, even at the aggregate level: We show that if the variance of idiosyncratic shocks is sufficiently high, the aggregate behaviour of the model will mimic a model with no labor mobility, even though in fact mobility will be high and the normative features of the equilibrium will be very different from that of a model with no mobility. This highlights the importance of second moments of moving costs (such as the variance of the idiosyncratic shocks) as well as the first moment, and points out an advantage of our structural approach over reduced-form econometric approaches.

The following section lays out the structure of the model. The subsequent section analyzes the solution to the planner's problem of the optimal rule for the allocation of labor, and finds the key Euler condition that characterizes optimality. The subsequent section shows that this optimal rule is implemented by the decentralized rational expectations equilibrium. The following section elaborates the most important properties of the equilibrium. Finally, we briefly discuss a special case of the model that offers a simple form for the equilibrium, facilitating empirical estimation.

[4] In a sense, this actually fits their data better than their own model, since they find that movers on average experience a loss in wages, which is the opposite of what one would expect if the point of moving was to find a higher wage.

[5] In addition, Topel (1986) requires the number of regions to be large so that asymptotic properties can be used to solve the equilibrium. Our model requires no such assumption.

2. The model.

Consider a model in which production may occur in any of n 'cells,' where a cell is taken to mean a particular industry in a particular place. For example, 'pharmaceuticals in New Jersey' might be one of the cells, as might 'pharmaceuticals in Delaware' or 'food service in New Jersey.' In each cell there are a large number of competitive employers, and the value of their aggregate output in any period t is given by $x_t^i = X^i(L_t^i, s_t) \geq 0$, where L_t^i denotes the labor used in cell i in period t, and s_t is a state variable that could capture the effects of policy (such as trade protection, which might raise the domestic price of the output), technology shocks, changes in world prices, and the like. Assume that s follows a first-order Markov process on some compact state space $S^s \subset \Re^k$ for some k, where the probability distribution for s_{t+1} conditional on s_t is given by a continuous density function $h(s_{t+1}; s_t)$.

Assume that X^i is strictly increasing, continuously differentiable and concave in its first argument, and also continuous in its second argument. Its first derivative with respect to labor, denoted X_1^i, is then a continuous, decreasing function of labor. We will assume that the price received by producers in a cell does not depend on the quantity produced in that cell,[6] so that X_1^i is the value marginal product of labor and thus the demand curve for labor in the cell.[7] Denote the total value of output by $x_t = X(L_t, s_t) \equiv \sum_i X^i(L_t^i, s_t)$.

The economy's workers form a continuum of measure $\overline{L}$. Each worker at any moment is located in one of the n cells. Denote the number of workers in cell i at the beginning of period t by L_t^i, and the allocation of workers by $L_t = [L_t^1, \ldots, L_t^n]$. If a worker, say, $\theta \in [0, \overline{L}]$, is in cell i at the beginning of t, she will produce in that cell, collect the market wage w_t^i for that cell, and then may move to any other cell.

If a worker moves from cell i to cell j, she incurs a cost $C^{ij} \geq 0$, which is the same for all workers and all periods, and is publicly known. This can include,

[6] For example, this would hold in the case of a small open economy in which the only trade impediments are tariffs, so that the domestic price of each good is equal to an exogenous world price plus a tariff rate. Another example would be if each location is small relative to the rest of the economy, each good is produced in a large number of locations within the country, and the market for each good is nationally integrated.

[7] This matters only for the property that the equilibrium can be represented as a distorted planner's optimum, which is useful for computation and for proof of some properties. The endogeneity of product prices is irrelevant for the market equilibrium conditions derived in Section 4 and for the estimation strategy outlined in Section 6.

for example, moving costs, if i and j are in different locations; training costs (tuition and time required for sector-specific schooling, for example) if i and j are in different industries; and a myriad of psychic costs as well that come from leaving a familiar location or occupation and moving to a new one. For example, in an economy with two sectors (textiles (T) and shoes (S)) and two regions (East (E) and West (W)), suppose that cells 1, 2, 3, and 4 are T-E (textiles-East), T-W, S-E and S-W respectively. In that case, C^{12}, C^{21}, C^{34}, and C^{43} are costs of moving between the regions, which include moving company services, realtors' fees, search costs for a new house, and the like. On the other hand, C^{13} and C^{24} are costs of moving out of the textile business and acquiring the human capital required to be an effective worker in the shoe business, which could involve night school or the time cost of making the right network connections for the new line of work.

In addition, if she is in cell i at the end of period t, the worker collects an idiosyncratic benefit $\varepsilon^i_{\theta,t}$ from being in that cell. These benefits are independently and identically distributed across individuals, cells, and dates, with density and cumulative distribution function f and $F : \Re \longmapsto \Re^+$ respectively, where $f(\varepsilon) > 0 \forall \varepsilon$. We normalize the average value $\int \varepsilon f(\varepsilon) d\varepsilon$ of the ε's to zero. One can think of these benefits as capturing anything in one's personal situation that may affect the direction or timing of labor market decisions independently of wages. For example, in the example of the previous paragraph, a worker in T-E may become terribly bored of the textile business and long for a change. This would correspond to a low value for ε^1 and ε^2. On the other hand, this person may fall in love with someone who lives in West, inducing high values for ε^2 and ε^4. Finally, the worker's family may have a member who is at the moment under the care of a trusted local doctor, or the children may be near the end of high school, and at the same time the worker has developed a good working rapport with her current employer. In that case, any move would be costly, and we have low values for ε^2, ε^3, and ε^4.

Thus, the full cost for worker θ of moving from i to j can be thought of as $\varepsilon^i_{\theta,t} - \varepsilon^j_{\theta,t} + C^{ij}$, the first two terms representing the idiosyncratic cost, and the last term the common cost. Note that the idiosyncratic cost can be *negative*, which is important, because that provides for gross labor flows in excess of net flows. Adopt the convention that $C^{ii} = 0$ for all i.

All agents have rational expectations and a common constant discount factor $\beta < 1$, and are risk neutral. Finally, we make the following boundedness assumption:

$$\int \varepsilon F^{n-1}(\varepsilon)f(\varepsilon)d\varepsilon < \infty. \tag{2.1}$$

This states that the expected value of the maximum ε for any worker on any date is finite.

3. The planner's problem.

It is useful to examine a hypothetical social planner's solution to the problem of allocating workers to cells in this framework. Note that we mean 'social planner' in a narrow sense. For example, it has already been made clear that the state variable s can include policy variables such as trade barriers, and these will all be treated as exogenous. Given this qualification, the social planner chooses an allocation rule, which can be summarized as a set of functions $D^{ij} : (\Re^n \times \Re^n \times S^s) \longmapsto [0,1]$, with the interpretation that $D^{ij}(\varepsilon; L, s)$ is the fraction of workers in cell i with idiosyncratic shocks $\varepsilon = (\varepsilon^1, \ldots, \varepsilon^n)$ who will be moved to cell j. Naturally, we must have

$$\sum_{j=1}^{n} D^{ij}(\varepsilon; L, s) = 1 \forall i \in \{1, \ldots n\}, \varepsilon, L \in \Re^n, \text{ and } s \in S^s. \tag{3.1}$$

The planner wishes to maximize:

$$E_{\{s_t\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \beta^t \sum_{i=1}^{n} \left[X^i(L_t^i, s_t) + L_t^i \int \cdots \int \left(\sum_{j=1}^{n} D^{ij}(\varepsilon; L_t, s_t)(\varepsilon^j - C^{ij}) \right) \prod_{j=1}^{n} (f(\varepsilon^j) d\varepsilon^j) \right], \tag{3.2}$$

subject to (3.1) and:

$$L_{t+1}^i = \sum_{k=1}^{n} L_t^k \int \cdots \int D^{ki}(\varepsilon; L_t, s_t) \prod_{j=1}^{n} (f(\varepsilon^j) d\varepsilon^j),$$

with L_0 and s_0 given, with respect to the functions D^{ij}. The first term in the square brackets of the objective function is simply the value of the output in cell

i, and the second term is the aggregate of idiosyncratic benefits ε^j, contingent on location decisions, and net of non-idiosyncratic moving costs C^{ij}. The constraint is simply the law of motion for the stock of workers in each cell: L^i_{t+1} equals the measure of period t cell i workers who remain there to period $t+1$, plus aggregate arrivals to i from other cells.

It will be convenient to denote by m^{ij}_t the fraction of workers in cell i who move to j in period t. Of course, this is equal to $\int \cdots \int D^{ij}(\varepsilon; L_t, s_t) \prod_{k=1}^n (f(\varepsilon^k) d\varepsilon^k)$. Given the full support assumption made for the ε's, it will be clear that it will be optimal to have $m^{ij}_t > 0 \forall i$, j, and t.

It is easy to demonstrate that an optimal allocation rule will always take a particular form. First, for any pair of cells, i and j, at each date and state, there is always a threshold, $\overline{\varepsilon}^{ij}$, such that no worker in i moves to j if her realization of $\varepsilon^i - \varepsilon^j$ is greater than $\overline{\varepsilon}^{ij}$, and no worker in i remains in i if her $\varepsilon^i - \varepsilon^j$ is less than $\overline{\varepsilon}^{ij}$. Thus, $\overline{\varepsilon}^{ij}$ may be interpreted as the marginal idiosyncratic moving cost for a mover from i to j. (Not surprisingly, later it will be seen that for an optimal allocation rule, $\overline{\varepsilon}^{ij}$ must also equal the marginal benefit to having one more worker moved from i to j, and thus it will reflect all available information about future labor demand in the two cells as well as the common moving costs, the C^{ij}'s.)

Proposition 3.1. *Consider an optimal allocation rule $\{D^{ij}\}_{i,j \in \{1,\ldots n\}}$. Fix i, $j \neq i$, t, L_t, and s_t, and suppose that at that state $m^{ij}_t, m^{ii}_t > 0$. For any number $\widetilde{\varepsilon}$, define:*

$$\chi(\widetilde{\varepsilon}) \equiv \int \left(\int_{-\infty}^{\infty} \int_{\varepsilon^j + \widetilde{\varepsilon}}^{\infty} D^{ij}(\varepsilon; L_t, s_t) f(\varepsilon^i) d\varepsilon^i f(\varepsilon^j) d\varepsilon^j \right) \prod_{k \neq i,j} \left(f(\varepsilon^k) d\varepsilon^k \right), \text{ and}$$

$$\xi(\widetilde{\varepsilon}) \equiv \int \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\varepsilon^j + \widetilde{\varepsilon}} D^{ii}(\varepsilon; L_t, s_t) f(\varepsilon^i) d\varepsilon^i f(\varepsilon^j) d\varepsilon^j \right) \prod_{k \neq i,j} \left(f(\varepsilon^k) d\varepsilon^k \right).$$

(In other words, for any number $\widetilde{\varepsilon}$, $\chi(\widetilde{\varepsilon})$ is the fraction of i workers who have $\varepsilon^i - \varepsilon^j > \widetilde{\varepsilon}$ and move to j; and $\xi(\widetilde{\varepsilon})$ is the fraction of i workers who have $\varepsilon^i - \varepsilon^j < \widetilde{\varepsilon}$ and remain in i.) Then there exists $\overline{\varepsilon}^{ij}$ such that $\chi(\overline{\varepsilon}^{ij}) = \xi(\overline{\varepsilon}^{ij}) = 0$.

We will adopt the convention that $\overline{\varepsilon}^{ii} = 0 \forall i$, and will denote the matrix of these thresholds as $\overline{\varepsilon} \equiv \{\overline{\varepsilon}^{ij}\}_{i,j \in (1,\ldots n)}$. An important note is that $\varepsilon^i - \varepsilon^j < \overline{\varepsilon}^{ij}$ does not *ensure* that the worker goes to j, because it is possible that she will choose a third option. That point is clarified by the following proposition, which shows how all of the $\overline{\varepsilon}^{ij}$ together fully determine the choices of each worker (to within a set of measure zero).

Proposition 3.2. *Let the conditions in the previous proposition hold, and suppose that we have chosen a set of $\bar{\varepsilon}^{ij}$ as described there. Then $D^{ij}(\varepsilon; L_t, s_t) = 1$ if and only if j solves:*

$$\max_{k\in\{1,\ldots n\}}\{\varepsilon^k + \bar{\varepsilon}^{ik}\}$$

(except possibly on a set of measure zero). Equivalently, $D^{ij}(\varepsilon; L_t, s_t) = 0$ if and only if j does not maximize $\{\varepsilon^k + \bar{\varepsilon}^{ik}\}$, except possibly on a set of measure zero.

This allows us to write the planner's problem in a simple way, as the choice of a function $\bar{\varepsilon}(L, s)$ giving the thresholds at each date and state. The realized current-period payoff to a given worker in cell i is equal to that worker's wage, w_t^i, plus $(\varepsilon^j - C^{ij})$, if that worker moves to cell j. Conditional on the $\bar{\varepsilon}^{ik}$'s and on ε_j, the probability that this worker does move to cell j is $\prod_{k\neq j} F(\varepsilon^j + \bar{\varepsilon}^{ij} - \bar{\varepsilon}^{ik})$. For this reason, the realized value of the objective function (3.2) will be:

$$E_{\{s_t\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(L_t, s_t, \bar{\varepsilon}(L_t, s_t)), \tag{3.3}$$

where

$$U(L, s, \bar{\varepsilon}) \equiv \sum_{i=1}^{n}\left[X^i(L^i, s) + L^i \sum_{j=1}^{n}\left(\int_{-\infty}^{\infty}(\varepsilon^j - C^{ij}) f(\varepsilon^j) \prod_{k\neq j} F(\varepsilon^j + \bar{\varepsilon}^{ij} - \bar{\varepsilon}^{ik}) d\varepsilon^j\right)\right]. \tag{3.4}$$

We can write the gross flows of workers out of sector i as a function of the $\bar{\varepsilon}^{ij}$'s:

$$m^{ij}(\bar{\varepsilon}^i) = \int_{-\infty}^{\infty} f(\varepsilon^j) \prod_{k\neq j} F(\varepsilon^j + \bar{\varepsilon}^{ij} - \bar{\varepsilon}^{ik}) d\varepsilon^j, \tag{3.5}$$

where $\bar{\varepsilon}^i = (\bar{\varepsilon}^{i1}, \ldots, \bar{\varepsilon}^{in})$. We can write $m^i(\bar{\varepsilon}^i) = (m^{i1}(\bar{\varepsilon}^i), \ldots m^{in}(\bar{\varepsilon}^i))$. This allows us to write the law of motion as a function of the $\bar{\varepsilon}^{ij}$'s:

$$L_{t+1}^i = m^{ii}(\bar{\varepsilon}^i) L_t^i + \sum_{k\neq i} m^{ki}(\bar{\varepsilon}^k) L_t^k, \text{ so} \tag{3.6}$$

$$L'_{t+1} = L'_t m(\bar{\varepsilon}), \tag{3.7}$$

where m denotes the full matrix of gross flows and a prime on a vector indicates the transpose.

The planner, then, maximizes (3.3) subject to (3.7), given L_0 and s_0.

Equation (3.5) defines all gross flows out of cell i as a function of $\overline{\varepsilon}^i$, with domain $\{\overline{\varepsilon}^i : \overline{\varepsilon}^{ij} \in \Re, \overline{\varepsilon}^{ii} = 0\}$ and range $\{m^i : m^{ij} > 0, \sum_j m^{ij} = 1\}$. The following presents a useful property of this function.

Proposition 3.3. *For any i, the function m^i is invertible.*

Thus, we can meaningfully write either the gross flows as a function of the $\overline{\varepsilon}^{ij}$'s (that is, $m^{ij}(\overline{\varepsilon}^{ij})$) or vice versa ($\overline{\varepsilon}^{ij}(m^{ij})$) without ambiguity. This result is useful partly because it is helpful in deriving the planner's first order condition. In addition, note that although the $\overline{\varepsilon}^{ij}$'s are useful from the point of view of theory, they are of course unobservable to an econometrician. However, in some cases the gross flows m^{ij} themselves *are* observable in conventional labor force surveys. This theorem gives us a way of inferring the values of the unobservable $\overline{\varepsilon}^{ij}$'s by studying the observable m^{ij}'s. This is a key to the econometric estimation of the model.

3.1. The planner's first order condition.

It is clear that the optimization problem presented above can be represented as a stationary dynamic programming problem, with Bellman equation:

$$V(L,s) = \max_{\overline{\varepsilon}}\{U(L,s,\overline{\varepsilon}) + \beta E_{\widetilde{s}}[V(\widetilde{L},\widetilde{s})|s]\}, \tag{3.8}$$

where $V : \Re^n \times S^s \mapsto \Re$ is the value function,[8] $\widetilde{L}$ and $\widetilde{s}$ are the next-period values of the labor allocation vector L and the state s, with $\widetilde{L}$ calculated from L and $\overline{\varepsilon}$ by (3.7), and where the expectation is taken with respect to the distribution of $\widetilde{s}$, conditional on s. Standard properties of dynamic programming problems will hold here; for example, the value function will be differentiable in L.[9] In addition:

[8] Of course, values for L in the solution will range only within the set $\{L \in \Re^n | L^i \geq 0; \sum_i L^i = \overline{L}\}$. It is useful, nonetheless, to define the optimization problem for all $L \in \Re^n$; for example, this makes the partial derivaties V_i, $i = 1, \ldots n$ meaningful.

[9] It is straightforward to verify that the conditions of Theorem 9.10 of Lucas and Stokey (1989, p. 266) are satisfied. Technically, to apply that theorem, we need to restrict the domain for L to a bounded set such as $S^L(L^*) \equiv \{L \in \Re^n | L^i \geq 0; \sum_i L^i \in [0, \overline{L}^*]\}$ for some $\overline{L}^* > \overline{L}$, to ensure boundedness of the objective function. However, this works for any value of $\overline{L}^*$, and so is not restrictive.

Proposition 3.4. *The value function is (i) non-negative; (ii) uniformly bounded on any compact subset of the domain; and (iii) concave in L.*

The first order condition with respect to the $\bar{\varepsilon}^{ij}$ terms can be obtained mechanically, and rearranged to yield the following.

Proposition 3.5. *In an optimal allocation, the condition:*

$$\bar{\varepsilon}^{ij} + C^{ij} = \beta E \left(\frac{\partial \widetilde{V}}{\partial \widetilde{L}^j} - \frac{\partial \widetilde{V}}{\partial \widetilde{L}^i} \right) \tag{3.9}$$

will hold at all times.

To interpret this condition, recall that $\bar{\varepsilon}^{ij}$ denotes the value of $\varepsilon^i - \varepsilon^j$ for the marginal mover from i to j, and is thus the marginal idiosyncratic cost of reallocating a worker from i to j. The left hand side of the equation is therefore the marginal cost of moving workers from cell i to cell j. The right hand side is the discounted marginal value of doing so.

In order to shed more light on the right-hand side of this condition, the envelope condition can be applied to the Bellman equation, yielding the following.

Proposition 3.6. *The marginal value of a worker in cell i in the optimal allocation satisfies:*

$$\frac{\partial V(L,s)}{\partial L^i} = X_1^i + \Omega(\bar{\varepsilon}^i) + \beta E \frac{\partial \tilde{V}}{\partial \widetilde{L}^i}, \tag{3.10}$$

where

$$\Omega(\bar{\varepsilon}^i) = \sum_{j=1}^{n} \int_{-\infty}^{\infty} (\varepsilon^j + \bar{\varepsilon}^{ij}) f(\varepsilon^j) \prod_{k \neq j} F(\varepsilon^j + \bar{\varepsilon}^{ij} - \bar{\varepsilon}^{ik}) d\varepsilon^j. \tag{3.11}$$

This equation has a natural interpretation. An increase in the number of workers in cell i has three effects. The first is the direct effect of increased production in cell i. The last is the benefit those workers generate in cell i if that is where they remain. The middle term, which is simply the average value of $\max_j\{\varepsilon^j + \bar{\varepsilon}_t^{ij}\}$ for all workers currently in cell i, is the additional benefit owing to the ability to reallocate these workers into other cells. The Ω function is thus a measure of the option value resulting from the ability to move workers from one cell to another.

Putting this together with (3.9) then yields the Euler equation for this optimization problem:

$$\begin{aligned}\bar{\varepsilon}_t^{ij} + C^{ij} &= \beta E_t[X_1^j(L_{t+1}^j, s_{t+1}) - X_1^i(L_{t+1}^i, s_{t+1}) + \Omega(\bar{\varepsilon}_{t+1}^j) - \Omega(\bar{\varepsilon}_{t+1}^i) \quad (3.12) \\ &+\beta\left(\frac{\partial V(L_{t+2}^j, s_{t+2})}{\partial L_{t+2}^j} - \frac{\partial V(L_{t+2}^i, s_{t+2})}{\partial L_{t+2}^i}\right)],\end{aligned}$$

which, then, using (3.9) again, can be rewritten in the useful form as follows.

Proposition 3.7. *In an optimal allocation, the condition:*

$$\bar{\varepsilon}_t^{ij} + C^{ij} = \beta E_t\left(X_1^j(L_{t+1}^j, s_{t+1}) - X_1^i(L_{t+1}^i, s_{t+1}) + \Omega(\bar{\varepsilon}_{t+1}^j) - \Omega(\bar{\varepsilon}_{t+1}^i) + \bar{\varepsilon}_{t+1}^{ij} + C^{ij}\right) \quad (3.13)$$

will hold at all times for all $i \neq j$.

This is the key condition for characterizing the optimal dynamic allocation of workers, and also for estimating the model econometrically. The economic meaning of this condition is fairly straightforward. Recall that the left-hand side of the equation is the marginal cost of moving workers from cell i to cell j at time t. The right-hand side is the marginal benefit of such a reallocation. As indicated in (3.9), that marginal benefit is equal to the difference in the expected discounted marginal value of a worker in the two cells next period. As indicated in (3.10), that difference has three parts: The direct effect of the difference in marginal value products in the two cells, the continuation value of leaving the reallocated workers in the new cell next period, and the additional value that comes from exercising the option of reallocating some portion of them next period. The direct effect is indicated by the first two terms on the right hand side of (3.13), and the difference in option values is indicated by the following two terms. The difference in continuation values is indicated by the final two terms, which are simply the marginal cost of moving a worker from i to j *next period*, which by *next period*'s first order condition (3.9) must be equal to the difference in the expected discounted marginal value of a worker in the two cells the *following* period.

4. Market Equilibrium.

Here we show that the optimal allocation rule analyzed above is also the equilibrium of a decentralized economy. Assume that all workers and employers take wages as given. In each cell i at each date t, the wage w_t^i will adjust to clear the market, so that $w_t^i = X_1^i(L_t^i, s_t)$ at all times. Assume that any worker who chooses to move from i to j will herself bear both the common moving cost, C^{ij}, and the idiosyncratic moving costs, $\varepsilon^i - \varepsilon^j$. All agents have rational expectations and a common constant discount factor $\beta < 1$, and are risk neutral.

An equilibrium then takes the form of a decision rule by which, in each period, each worker will decide whether to stay in her cell or move to another, based on the current allocation vector L of labor across sectors, the current aggregate state s, and that worker's own vector ε of shocks. In the aggregate, this decision rule generates a law of motion for the evolution of labor allocation and, by the labor market clearing condition just mentioned, for the wage in each sector. Given this behaviour for wages, the decision rule must be optimal for each worker, in the sense of maximizing her expected present discounted value of wages plus idiosyncratic benefits net of moving costs.

Let the maximized value to each worker of being in sector i when the labor allocation is L and the state is s be denoted by $\widehat{v}^i(L, s, \varepsilon)$, which, of course, depends on the worker's realized idiosyncratic shocks. Denote by $v^i(L, s)$ the average of $\widehat{v}^i(L, s, \varepsilon)$ across all workers in i, or in other words, the expectation of $\widehat{v}^i(L, s, \varepsilon)$ with respect to the vector ε. Thus, $v^i(L, s)$ can also be interpreted as the expected value of being in cell i, conditional on L and s, but before the worker learns her value of ε. In contrast to the previous section, define $\overline{\varepsilon}_t^{ij}$ by:

$$\overline{\varepsilon}_t^{ij} \equiv \beta E_t[v^j(L_{t+1}, s_{t+1}) - v^i(L_{t+1}, s_{t+1})] - C^{ij}. \tag{4.1}$$

We can think of this definition of $\overline{\varepsilon}_t^{ij}$ as indicating the common net benefit to moving from cell i to cell j, which each i worker will then weigh against the idiosyncratic costs of moving. This definition of $\overline{\varepsilon}_t^{ij}$ will be seen to be equivalent to the definition of the previous section.

We can write a typical i-worker's optimization problem as follows:

$$\begin{aligned} \widehat{v}^i(L_t, s_t, \varepsilon_t) &= w_t^i + \max_j\{\varepsilon_t^j - C^{ij} + \beta E_t[v^j(L_{t+1}, s_{t+1})]\} \\ &= w_t^i + \beta E_t[v^i(L_{t+1}, s_{t+1})] + \max_j\{\varepsilon_t^j + \overline{\varepsilon}_t^{ij}\}. \end{aligned} \tag{4.2}$$

Taking the expectation of (4.2) with respect to the ε vector then yields the i-worker's Bellman equation:

$$\upsilon^i(L_t, s_t) = w_t^i + \beta E_t[\upsilon^i(L_{t+1}, s_{t+1})] + \Omega(\overline{\varepsilon}_t^i), \tag{4.3}$$

where Ω is as defined in (3.11). Using (4.3), we can rewrite (4.1) as:

$$\begin{aligned} C^{ij} + \overline{\varepsilon}_t^{ij} &= \beta E_t[\upsilon^j(L_{t+1}, s_{t+1}) - \upsilon^i(L_{t+1}, s_{t+1})] \\ &= \beta E_t[w_{t+1}^j - w_{t+1}^i + \beta E_{t+1}[\upsilon^j(L_{t+2}, s_{t+2}) - \upsilon^i(L_{t+2}, s_{t+2})] \\ &\quad + \Omega(\overline{\varepsilon}_{t+1}^j) - \Omega(\overline{\varepsilon}_{t+1}^i)] \\ &= \beta E_t[w_{t+1}^j - w_{t+1}^i + C^{ij} + \overline{\varepsilon}_{t+1}^{ij} + \Omega(\overline{\varepsilon}_{t+1}^j) - \Omega(\overline{\varepsilon}_{t+1}^i)] \end{aligned} \tag{4.4}$$

Now we can verify that the allocation rule that solves the planner's problem is an equilibrium. First, for every t, L_t, and s_t, set $\upsilon^i(L_t, s_t)$ equal to $\frac{\partial V(L_t, s_t)}{\partial L_t^i}$ from the solution to the planner's problem. Then, at every state and date, the $\overline{\varepsilon}_t^{ij}$ matrix defined by (4.1) is the same as that induced by the solution to the planner's problem, as shown by (3.9). Further, from (3.10), together with the labor market clearing condition that the wage in each cell will be equal to the marginal value product of labor, it is clear that the υ^i functions constructed in this way satisfy the worker's Bellman equation (4.3). Of course, then (3.13), the condition that characterizes the planner's optimum, is equivalent to (4.4), which is the equilibrium condition characterizing the decentralized equilibrium. Thus, the rule that solves the planner's problem also describes a market equilibrium.

Henceforth, we can refer to the planner's optimum and the equilibrium interchangeably. Since the planner's problem is well-behaved, we thus have a proof of the existence of competitive equilibrium and a method for computing it as well.

Two remarks on this are called for. First, this is an optimum allocation in an extremely restricted sense. In particular, it should be pointed out once again that the optimization problem analyzed here takes trade policy (for example) as given, and looks only at the allocation of workers conditional on it. In addition, the objective function set up in the previous section did not make any allowance at all for distributional values. However, a large part of the interest in this problem springs from distribution values in practice; the point is, precisely, to analyze who the gainers and losers will be from a given change in trade policy, for example, and how badly the latter will be hurt. Thus, the actual objective function for policy analysis will generally be different from that studied above. Second, Dixit and Rob (1994) have pointed out that in the presence of risk aversion and missing insurance markets, the market equilibrium of a labor adjustment process with rational expectations will not generally be even constrained efficient. Those elements have not been incorporated into this model, but it would be of interest to incorporate them into a later version.

The point of the equivalence of the optimal rule and the decentralized equilibrium is not to make an argument for nonintervention in the adjustment process, but to facilitate more convenient computation and analysis of the decentralized equilibrium. We now turn to that task.

5. Properties of the equilibrium.

A number of key properties of the adjustment process can now be seen immediately.

(i) *Continual reallocation of workers.* Consider a special case of the model in which the state variable s is a constant. Then one can analyze steady states of the model, which can be calculated in the following way. For any matrix of $\bar{\varepsilon}^{ij}$'s, one can compute a matrix of gross flow rates from (3.5), and holding those flow rates constant one can compute steady-state values of the labor allocation vector L from (3.7). All of this information can then be used to calculate the right hand side of (3.13) for any $i \neq j$. Subtracting C^{ij}, one can then compare the result with $\bar{\varepsilon}^{ij}$. A fixed point of this process is then a steady state. Since this computation induces a continuous function, a steady state must exist. Label the steady state value of the $\bar{\varepsilon}^{ij}$ matrix so computed $\bar{\varepsilon}^{*}$, the associated matrix of gross flows m^{*}, and the associated steady state labor allocation vector L^{*}.

The point is that even at this steady state, there will still be a constant reallocation of workers. This is because the integrals in (3.5) will always have positive values. The reason is that the workers experience idiosyncratic shocks constantly, and each one will wish to change jobs or to move periodically for personal reasons. Thus, the model has no trouble accommodating the empirical fact that gross flows are much larger than net flows.

(ii) *Gradual adjustment.* Empirically, labor adjustment tends to occur gradually (for example, see evidence summarized by Rappaport (2000) on the intertemporal persistence of labor flows across US locations). It is easy to see that this is a feature of the present model as well. Indeed, if the economy is in a steady state and a shock occurs that changes the steady state allocation, the economy will not reach the new economy in any finite time. To see this, consider once again the special case in which s is a constant. Suppose that the economy's steady state allocation vector is L^*, with an associated steady state value $\overline{\varepsilon}^*$ of the $\overline{\varepsilon}^{ij}$ matrix and associated matrix of gross flows m^*. Denote the labor allocation matrix at time t by L_t, and suppose that $L_0 \neq L^*$. Suppose that at time T, $0 < T < \infty$, the economy is in the steady state. Then at time $t = T - 1$, the right hand side of (3.13) will take its steady state values, so the values of $\overline{\varepsilon}^{ij}$ on the left hand side must be equal to the corresponding elements of $\overline{\varepsilon}^*$. But then (3.5), the matrix of gross flows m^{ij}_{T-1} at time $T-1$ must equal the values in m^*. But then working backward from the law of motion (3.6), we find that L_{T-1} must be equal to L^*.[10] Continuing in this logic, we find that $L_0 = L^*$, which is a contradiction.

Thus, the economy can move only gradually toward the steady state if it is not already in it. The reason is again the idiosyncratic shocks. Suppose that a given sector has enjoyed protection from imports for many years but suddenly the protection is taken away, and the change is expected to be permanent. The demand for labor in the sector drops, and the result is a reduction in the wages it pays; workers begin to reallocate themselves to other sectors, but each period a fraction of the workers waits because for those workers the cost of moving is high, and it is in their interest to wait in hopes of a lower draw for their moving costs in the near future.

(iii) *Anticipatory movement of workers.* In general, in this model if a change in labor demand in some cell is foreseen, that will result in a movement of workers before the fact. This can be seen most easily in a two-cell version of the model. Suppose that cell 1 is an export sector and cell 2 is an import-competing sector, which is protected by a tariff. At time 0, the government announces that it will eliminate the tariff beginning in period $T > 0$. There are no other changes in the economy at any time. This can be incorporated into the model by letting $s_t = t \forall t$, and by letting $X^2(\cdot, s)$ have one functional form when $s \geq T$ and a different one when $s < T$. The function is shifted down and flatter when $s \geq T$ compared with when $s < T$, since the tariff elevates the domestic price of cell 2's output, and hence the marginal value product of cell-2 labor. Let L^*, $\overline{\varepsilon}^*$ and m^* denote the steady state values for the economy with the tariff in place and

[10] Given L^* and m^*, the equation $(m^*)^{11} L^1_{T-1} + (m^*)^{21}(\overline{L} - L^1_{T-1}) = (L^*)^1$ has a unique solution for L^1_{T-1} provided that $(m^*)^{11} \neq (m^*)^{21}$. Given that $m^{11} = \Pr[\varepsilon^1 > \varepsilon^2 + \overline{\varepsilon}^{12}] = \Pr[\varepsilon^1 > \varepsilon^2 + \beta[\widetilde{V}_2 - \widetilde{V}_1] - C^{12}]$ and $m^{21} = \Pr[\varepsilon^1 + \overline{\varepsilon}^{21} > \varepsilon^2] = \Pr[\varepsilon^1 > \varepsilon^2 + \beta[\widetilde{V}_2 - \widetilde{V}_1] + C^{21}]$, $m^{11} > m^{21}$ provided that either C^{12} or $C^{21} > 0$.

expected to remain permanently (call this the 'tariff-affected steady state'), and suppose that $L_0 = L^*$. It can be seen quickly that no matter how large T is, the adjustment begins immediately, in the sense that because of the announcement the gross flows even in period 0 are already different from m^*.

To make the argument, it helps to consider two different stationary models, each with S^S a singleton, so that we can drop s as an argument in the value function. The first model (the 'starred' model) is one in which there is a tariff in place permanently, and the second model (the 'double-starred' model) is one in which there is never any tariff. The values L^*, $\widetilde{\varepsilon}^*$, and m^*, then, describe the steady-state of the 'starred' model. Denote the revenue functions for cell 2 for the two models by X^{2*} and X^{2**}, respectively. Apart from these revenue functions, the two models are identical. Denote the value functions by V^* and V^{**} respectively. The following property is easy to verify.

Proposition 5.1. *Assume that $X_1^{2*}(L^2) > X_1^{2**}(L^2)$ for all $L^2 > 0$. Then $dV^*(\overline{L} - L^2, L^2)/dL^2 > dV^{**}(\overline{L} - L^2, L^2)/dL^2$ for all $L^2 \in (0, \overline{L}]$ (call this the 'strong derivative property').*

Clearly, $V(L_t, s_t) = V^{**}(L_t)$ for $t \geq T$. The first-order condition at $t = T - 1$ is:

$$\overline{\varepsilon}_{T-1}^{12} + C^{12} = \beta[dV^{**}(\overline{L} - \widetilde{L}^2, \widetilde{L}^2)/d\widetilde{L}^2],$$

where a tilde denotes a next-period value. Given that $\overline{\varepsilon}_{t+1}^{21} = -\overline{\varepsilon}_{t+1}^{12} - C^{12} - C^{21}$ at all times (see (3.9)), we can think of $\overline{\varepsilon}_{t+1}^{21}$ as a decreasing function of $\overline{\varepsilon}_{t+1}^{12}$. Thus, an increase in $\overline{\varepsilon}^{12}$ will increase m^{12} and decrease m^{21}, increasing the next-period value of L^2. By the concavity of V^{**}, this will decrease the value of the right-hand side of the first-order condition. Thus, the right-hand side of the condition is a downward-sloping curve in $\overline{\varepsilon}^{12}$, while the left hand side is an upward-sloping line in $\overline{\varepsilon}^{12}$. As a result, for a given value of L^2, anything that shifts the right-hand side of the first-order condition down will result in a lower value of $\overline{\varepsilon}^{12}$. Therefore, by Proposition (5.1), the solution to the first-order condition at time $T - 1$ will yield a lower value of $\overline{\varepsilon}^{12}$, and thus a higher value of $\overline{\varepsilon}^{21}$, along with a lower value for the right-hand side of the first-order condition, than would be chosen for the same value of L_{T-1}^2 in the 'starred' model. Using the envelope condition (3.10), this implies that

$$dV(\overline{L} - L_{T-1}^2, L_{T-1}^2; s_{T-1})/dL_{T-1}^2 < dV_2^*(\overline{L} - L_{T-1}^2, L_{T-1}^2)/dL_{T-1}^2$$

for any $L^2_{T-1} \in (0, \overline{L}]$. Applying this same logic recursively back to $t = 0$, we conclude that the value of $\overline{\varepsilon}^{12}_0$ (and hence the value of m^{12}_0) that is chosen is below the value m^{*12} that would have been chosen in the steady state of the 'starred' model. But that demonstrates the point: The response to the future announced policy begins at the moment it is announced.

The interpretation of this result has to do once again with idiosyncratic shocks. Even if wages are currently equal in the two sectors, if a worker knows that an event will occur shortly in the future that will depress wages in sector 2 for a long time afterward, and if that worker happens to have low moving costs at the moment, understanding that her moving costs may not be so low later on, she may simply jump at the opportunity to move now. For example, a worker who has been separated from one firm in the sector that will experience the shock, instead of looking for employment with another firm in the same sector, may simply move to the other now that it is as easy to find a job there as in the worker's current sector.

It should be noted that anticipatory movements of labor are also a feature of Mussa-type models, as studied in detail by Dehejia (1997). However, in those models, the anticipatory behavior is a result of the existence of a retraining sector with rising marginal costs, while in the current model it arises purely from the presence of time-varying idiosyncratic moving costs. Anticipatory reorientation of an economy associated with a forthcoming change in trade policy is an important phenomenon empirically, as documented for the case of accessions to trade blocs by Freund and McLaren (1999). This mechanism provides an additional potential source for it.

(iv) *Anticipatory changes in wages.* This is an immediate corollary to the point just made. In the example discussed above, if workers begin to leave sector 2 immediately as soon as the planned future liberalization is announced, then clearly wages in sector 2 will begin to rise right away and wages in sector 1 will begin to fall right away. Of course, sector 2 wages will then drop abruptly at the date of the actual liberalization, and continue to adjust after that.

This is important for a number of reasons. First, in doing empirical work on the relationship between tariffs and wages, the issue of timing could be extremely important. Simply looking at a pair of snapshots taken before and after a liberalization, for example, could miss a large part of the actual movement in wages; further, in the simple story just told, if the pre-liberalization data were collected very shortly before the liberalization, the empirical results would overstate the downward effect of the liberalization on wages in the affected sector.

Second, these anticipatory effects on wages can provide a motive for gradualism in trade policy. If the government wishes to compensate the workers harmed by a liberalization but cannot do so through lump-sum transfers, announcing the policy change in advance and allowing these adjustment mechanisms to do their work can in principle be an effective way of doing so. This is a point made by Dehejia (1997) in the context of a Mussa-type model.

(v) *Persistent wage differentials in long-run equilibrium.* A feature of the model that is not obvious is that it generally predicts wage differentials across cells even in the steady state.

Consider, once again, a version with two cells and with s constant. Suppose that $C^{12} = C^{21}$, and suppose that there is a steady state in which $w^2 \geq w^1$. Observe that if in that steady state $L^1 > L^2$, then we must have $m^{21} > m^{12}$. From (3.5), this implies that $\bar{\varepsilon}^{21} > \bar{\varepsilon}^{12}$. Recalling that $\Omega(\bar{\varepsilon}^i) = E_\varepsilon[\max_j\{\varepsilon^j + \bar{\varepsilon}^{ij}\}]$, this implies that $\Omega(\bar{\varepsilon}^2) > \Omega(\bar{\varepsilon}^1)$. From (4.3) applied recursively, that means that $\upsilon^2 > \upsilon^1$. But from (4.1), this implies that $\bar{\varepsilon}^{21} < \bar{\varepsilon}^{12}$, a contradiction. Thus, in order to have $L^1 > L^2$ in the steady state, we must also have $w^1 > w^2$. *Thus, in the steady state a sector will have a higher wage than the other if and only if it has more workers than the other.* This conclusion contrasts sharply with the behavior of a Mussa-type model, in which factor returns are equalized across sectors in the long run (see Mussa (1978)).

The explanation is as follows. Suppose that both cells had the same wage in the steady state, but cell 1 was ten times the size of cell 2. In that case, workers would be indifferent between the two cells apart from idiosyncratic effects. In each period, a certain fraction of the workers in either cell would realize negative moving costs, which could be interpreted as boredom with the current job or location or a desire to move to the other cell to realize some personal opportunity. With the wages identical, an identical *fraction* of the workers in each cell would wish to change sectors in each period. However, this would imply a much larger *number* of workers moving from 1 to 2 than vice versa. The result would be net migration toward 2, which would push down the wage in cell 2 and pull up the wage in cell 1. The wage differential thus created would then tend to slow down migration out of 1 and speed up migration out of 2, and this process would continue until the aggregate number of workers moving in each direction would be equal.

These effects, which might be called 'frictional' wage differentials, thus provide a new reason for persistent intersectoral or geographic wage differences, quite independent of compensating differentials, efficiency wages and union effects, which have been emphasized in the labor economics literature. It should also be emphasized that these effects occur even if the average moving costs C^{ij} are all equal to zero. The persistent wage differentials are induced entirely by the variance in idiosyncratic effects.

(vi) Limiting behaviour as idiosyncratic shocks become important. Finally, there is a sense in which the aggregate behaviour of the model when idiosyncratic shocks are very important mimics the aggregate behaviour of a static model with no mobility at all. This underlines how crucial it is to take account of gross flows, as is being done here, and to estimate the structural parameters of the mobility costs, because using a reduced-form econometric approach could produce normative conclusions that would be seriously in error.

To make this point, consider a class of distributions for the ε^i's indexed by $\delta > 0$ in the following way. For a particular distribution function G_1 and associated density g_1, the distribution function G_δ and density g_δ are defined by $G_\delta(\varepsilon) = G_1(\varepsilon/\delta)$ and $g_\delta(\varepsilon) = g_1(\varepsilon/\delta)/\delta$. Thus, G_δ is a radial mean-preserving spread of G_1 for $\delta > 1$; the probability that $\varepsilon \leq y$ with the distribution G_1 is equal to the probability that $\varepsilon \leq \delta y$ with the distribution G_δ. With this family of distributions, if δ is very small, then idiosyncratic effects are trivial most of the time, but as δ becomes large, idiosyncratic effects become more important and can eventually dwarf wages in their effect on workers' decisions. The asymptotic effects of increases in δ are summarized in the following.

Proposition 5.2. *When the distribution of idiosyncratic shocks is given by the family G_δ, as $\delta \to \infty$ the matrix of gross flows m^{ij} converges uniformly in equilibrium over the whole state space to a matrix each of whose components is equal to $1/n$.*

Thus, if δ is very large, regardless of the labor demand shocks, workers would always be approximately evenly distributed across the cells of the economy. In this extreme case, which is certainly not realistic but a useful thought experiment to make a point, the number of workers in each cell would be completely insensitive to, for example, the elimination of tariffs, and all of the adjustment would occur in the form of changes in wages. Aggregate data would suggest that each industry has in effect a captive labor force, and the cost of the elimination of a tariff on textiles, for example, would be borne entirely by workers in the textile sector, while all other workers would enjoy a net benefit through lower textile prices. However, this would be quite wrong. In such an economy, far from being captive, workers would be very footloose, and a typical textile worker would face only a $1/n$

chance of continuing in the textile sector next period. Therefore, particularly if n is large, the cost borne by the textile workers would be very low; for most of such a worker's future career, she would be in other sectors, enjoying the benefit of lower prices. It may in fact be a Pareto-improving liberalization, while the reduced-form approach would mistakenly conclude that one sector of workers would be badly hurt and would bitterly oppose the liberalization. Thus, a focus on gross flows in equilibrium, and attention to the *variance* of mobility costs as well as their means, are, in principle, crucial to getting the normative conclusions right.

6. A special case, and empirical implementation.

The model takes a particularly tractable form when a judicious choice of functional form is made. Assume that the ε_t^i are generated from an extreme-value distribution with parameters $(-\gamma\nu, \nu)$, which implies:[11]

$$\begin{aligned} E[\varepsilon_t^i] &= 0 \quad \forall i, t \\ Var[\varepsilon_t^i] &= \frac{\pi^2\nu^2}{6} \quad \forall i, t. \end{aligned}$$

Note that while we make the natural assumption that the ε's be mean-zero, we do not impose any restrictions on the variance, leaving ν (which is positively related to the variance) as a free parameter to be estimated.

It can easily be shown that, with this assumption:

$$\bar{\varepsilon}_t^{ij} \equiv \beta E_t[V_{t+1}^j - V_{t+1}^i] - C^{ij} = \nu[\ln m_t^{ij} - \ln m_t^{ii}] \tag{6.1}$$

and:

$$\Omega(\bar{\varepsilon}_t^i) = -\nu \ln m_t^{ii}. \tag{6.2}$$

[11] The cumulative distribution, mean, and variance for an extreme-value distribution with parameters (α, ν) are given by:

$$\begin{aligned} F(\varepsilon) &= \exp\left\{-e^{-(\varepsilon-\alpha)/\nu}\right\} \\ E(\varepsilon) &= \alpha + \gamma\nu \\ Var(\varepsilon) &= \frac{\pi^2\nu^2}{6} \end{aligned}$$

For further properties of the extreme-value distribution, see Patel, Kapadia, and Owen (1976).

Both these expressions make intuitive sense. The first says that the greater the expected net (of moving costs) benefits of moving to j, the larger should be the observed ratio of movers (from i to j) to stayers. Moreover, holding constant the (average) expected net benefits of moving, a higher variance of the idiosyncratic cost shocks lowers the compensating migratory flow if the average net benefit is positive and raises it if they are negative.

The second expression says that the greater the probability of remaining in cell i, the lower the value of having the option to move from cell i.[12] Moreover, as one might expect, when the variance of the idiosyncratic component of moving costs increases, so too does the value of having the option to move.

Substituting from (6.1) and (6.2) into (3.13) we get:

$$\begin{aligned} C^{ij} + \nu[\ln m_t^{ij} - \ln m_t^{ii}] &= \beta E_t[w_{t+1}^j - w_{t+1}^i + C^{ij} + \nu[\ln m_{t+1}^{ij} - \ln m_{t+1}^{ii}] \\ &\quad + \nu[\ln m_{t+1}^{ii} - \ln m_{t+1}^{jj}]] \end{aligned}$$

This expression can be simplified and rewritten as the following conditional moment restriction:

$$E_t\left[\frac{\beta}{\nu}(w_{t+1}^j - w_{t+1}^i) + \beta(\ln m_{t+1}^{ij} - \ln m_{t+1}^{jj}) - \frac{(1-\beta)}{\nu}C^{ij} - (\ln m_t^{ij} - \ln m_t^{ii})\right] = 0 \tag{6.3}$$

This has the virtue that it can be estimated with data on gross flows and wages, using standard Generalized Method of Moment techniques. This is an ongoing project.

7. Conclusion.

This paper has articulated an equilibrium model of labor adjustment to external shocks, which has been designed to be useful for trade policy analysis and to be empirically estimable. The key features are an infinite horizon in which all workers have rational expectations; the possibility of shocks to labor demand in a sector (as caused, for example, by a change in trade policy) or in a geographic location; publicly observable costs of moving or of changing sectors; and time-varying, idiosyncratic private costs as well. We have shown that the equilibrium solves a particular social planner's dynamic programming problem, which facilitates analysis of the equilibrium. In addition, the equilibrium exhibits gross flows in excess

[12] Note that $0 < m_t^{ii} < 1$, so $\Omega(\overline{\varepsilon}_t^i) = -\nu \ln m_t^{ii} > 0$.

of net flows (and indeed, constant movement of workers even in a steady state), which is an important feature of empirical labor adjustment; gradual adjustment to a shock; anticipatory adjustment to an announced policy change; and persistent 'frictional' wage differentials across geographic locations or sectors, which will exist even if the average moving costs are zero, and which provide a new and independent theoretical rationale for wage differentials in long-run equilibrium. We have also shown, through a simple thought experiment, why the idiosyncratic shocks we highlight and the variance of those shocks are potentially so important to the normative conclusions in applied work.

Finally, it is shown that the key equilibrium condition takes a particularly simple form when the functional forms are chosen in a particular way, making the econometric estimation of the parameters of the model feasible with data on gross flows and wages over time for a particular economy. This is the subject of ongoing work.

8. Appendix.

Proof of Proposition (3.1). Clearly $\chi(\widetilde{\varepsilon})$ is decreasing and continuous, with $\chi(\widetilde{\varepsilon}) \to 0$ as $\widetilde{\varepsilon} \to \infty$ and $\chi(\widetilde{\varepsilon}) \to m_t^{ij}$ as $\widetilde{\varepsilon} \to -\infty$. Clearly $\xi(\widetilde{\varepsilon})$ is increasing and continuous, with $\xi(\widetilde{\varepsilon}) \to m_t^{ii}$ as $\widetilde{\varepsilon} \to \infty$ and $\chi(\widetilde{\varepsilon}) \to 0$ as $\widetilde{\varepsilon} \to -\infty$. Thus, we can find an $\overline{\varepsilon}^*$ such that $\chi(\overline{\varepsilon}^*) = \xi(\overline{\varepsilon}^*)$. If $\chi(\overline{\varepsilon}^*) = 0$, we are done. If not, then we have a positive mass of i workers who have $\varepsilon^i - \varepsilon^j < \overline{\varepsilon}^*$ and who remain in i, and an equal mass of i workers who have $\varepsilon^i - \varepsilon^j > \overline{\varepsilon}^*$ and who move to j. Clearly, if we simply reversed their roles, making the movers stay and the stayers move, the next-period allocation of labor would be unchanged, and the total surplus would be higher. Therefore, the original allocation rule could not have been optimal. ■

Proof of Proposition (3.2). Consider an optimal allocation. Suppose that for some i, j, k, L_t, s_t, and some set $A(1) \subseteq \Re^n$ with positive probability measure, $\varepsilon^j + \overline{\varepsilon}^{ij} > \varepsilon^k + \overline{\varepsilon}^{ik}$ and yet $D^{ik}(\varepsilon; L_t, s_t) > 0 \forall \varepsilon \in A(1)$. Without loss of generality, assume that for all $\varepsilon \in A(1)$, $\varepsilon^j + \overline{\varepsilon}^{ij} - (\varepsilon^k + \overline{\varepsilon}^{ik}) \geq \widetilde{\varepsilon} > 0$. For any positive N, consider the ball of radius $1/N$ around the point $\varepsilon = (-\overline{\varepsilon}^{i1}, -\overline{\varepsilon}^{i2}, \ldots, -\overline{\varepsilon}^{in})$, and note that within such a ball will be points for which the expression $\varepsilon^i - \varepsilon^{i'} - \overline{\varepsilon}^{ii'}$ is negative for all i', points for which it is positive for all i', and points with every other possible combination of signs (note that at the center of the ball $\varepsilon^i - \varepsilon^{i'} -$

$\overline{\varepsilon}^{ii'} = 0 \forall i')$. For $N = 1, \ldots, \infty$, define a subset of such a ball, $B(N) \subseteq \Re^n$, by $B(N) = \{\varepsilon : \varepsilon^i - \varepsilon^{i'} > \overline{\varepsilon}^{ii'} \forall i' \neq j; \varepsilon^i - \varepsilon^j < \overline{\varepsilon}^{ij}; and \max_{i'} \left|\varepsilon^{i'} + \overline{\varepsilon}^{ii'}\right| < 1/N\}$. (Note that at the center of the ball, $\varepsilon^{i'} + \overline{\varepsilon}^{ii'} = 0 \forall i'$.) By the previous proposition, $D^{ij} = 1$ everywhere on $B(N)$ for all N. Define a sequence $A(N)$ of subsets of $A(1)$, where for each N the probability measure $p(N) \equiv \int_{A(N)} D^{ik}(\varepsilon; L_t, s_t) \prod_{k=1}^{n} (f(\varepsilon^k) d\varepsilon^k)$ of workers in $A(N)$ who go to k is equal to the smaller of $p(1)$ and the measure of $B(N)$. For large enough N, we will have $\varepsilon^j + \overline{\varepsilon}^{ij} - (\varepsilon^k + \overline{\varepsilon}^{ik}) < \widetilde{\varepsilon}$ for all $\varepsilon \in B(N)$, and a measure of workers in $A(N)$ going to k that is equal to the measure of workers in $B(N)$ who go to j. But then for every worker in $A(N)$, $\varepsilon^j - \varepsilon^k \geq \overline{\varepsilon}^{ik} - \overline{\varepsilon}^{ij} + \widetilde{\varepsilon}$, and the worker moves to k; while for every worker in $B(N)$, $\varepsilon^j - \varepsilon^k < \overline{\varepsilon}^{ik} - \overline{\varepsilon}^{ij} + \widetilde{\varepsilon}$, and the worker moves to j. Clearly, if for $\varepsilon \in A(N)$, we simply reduced $D^{ik}(\varepsilon; L_t, s_t)$ to 0 and increased $D^{ij}(\varepsilon; L_t, s_t)$ by $D^{ik}(\varepsilon; L_t, s_t)$; and if for $\varepsilon \in B(N)$, we reduced $D^{ij}(\varepsilon; L_t, s_t)$ to 0 and increased $D^{ik}(\varepsilon; L_t, s_t)$ to 1; then the total number of workers going to each cell would be unchanged. However, a positive mass of workers in $A(N)$ and in $B(N)$ will have reversed their roles; $B(N)$ workers with lower values of $\varepsilon^j - \varepsilon^k$ now move to k and the $A(N)$ workers with higher values of $\varepsilon^j - \varepsilon^k$ move to j. Thus, the next-period allocation of labor would be unchanged, and the total surplus would be higher. Therefore, the original allocation rule could not have been optimal. ■

Proof of Proposition (3.3). Recall the gross flow function defined by (3.5). It is convenient to define a truncated version of this function. First, let x_{-k} denote the vector made by deleting the k^{th} element of x (if x has fewer than k elements, $x_{-k} = x$). After deleting one or more elements of a vector, continue to index the remaining elements in the same way, so, for example, if $x \in \Re^n$ and $n > i$, then $(x_{-i})^n = x^n$. In addition, for any vector x, let $x^{[k]}$ denote the vector made up of its first k elements; let $x^{-[k]}$ denote the vector made up of all of its elements after the k^{th}.

Then, for any i, define $\widetilde{m}^i : \Re^{n-1} \to \{x \in (0,1)^{n-1} : \sum_j x^j < 1\}$, with $\widetilde{m}^i(\overline{\varepsilon}^i_{-i}) = (m^i(\overline{\varepsilon}^i))_{-i}$. Thus, $\widetilde{m}^i$ defines the gross flows out of i, but not the residual category of i workers who stay in i, and it defines them as a function of $\overline{\varepsilon}^i_{-i}$.

We now derive some information about the derivatives of $\widetilde{m}^{ij}$. They are as follows:

$$\frac{\partial \widetilde{m}^{ij}(\overline{\varepsilon}^i_{-i})}{\partial \overline{\varepsilon}^{ii'}} = -\int_{-\infty}^{\infty} f(\varepsilon^j) f(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ii'}) \prod_{k \neq j, i'} F(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}) d\varepsilon^j < 0$$

if $i' \neq j$, and

$$\int_{-\infty}^{\infty} f(\varepsilon^{i'}) \sum_{k \neq i'} f(\varepsilon^{i'} + \overline{\varepsilon}^{ii'} - \overline{\varepsilon}^{ik}) \prod_{l \neq i', k} F(\varepsilon^{i'} + \overline{\varepsilon}^{ii'} - \overline{\varepsilon}^{il}) d\varepsilon^{i'} > 0 \tag{8.1}$$

if $i' = j$.

Note that if $i \neq i'$,

$$\begin{aligned} \sum_{j \neq i} \frac{\partial \widetilde{m}^{ij}(\overline{\varepsilon}^{i}_{-i})}{\partial \overline{\varepsilon}^{ii'}} &= -\frac{\partial m^{ii}(\overline{\varepsilon}^{i}_{-i})}{\partial \overline{\varepsilon}^{ii'}} \\ &= \int_{-\infty}^{\infty} f(\varepsilon^{i}) f(\varepsilon^{i} - \overline{\varepsilon}^{ii'}) \prod_{k \neq i, i'} F(\varepsilon^{i} - \overline{\varepsilon}^{ik}) d\varepsilon^{i} \\ &> 0. \end{aligned}$$

Thus, the matrix of derivatives

$$\nabla \widetilde{m}^{i} \equiv \left(\frac{\partial \widetilde{m}^{ij}(\overline{\varepsilon}^{i}_{-i})}{\partial \overline{\varepsilon}^{ii'}} \right)_{j, i' \neq i},$$

which is the Jacobian of the $\widetilde{m}^{i}$ function, is a dominant diagonal matrix with positive elements on the main diagonal and negative elements off the main diagonal. This implies that it has an inverse (see Theorem 1 in McKenzie (1960)), and that the inverse has only positive elements (see Theorem 4 in McKenzie (1960)). This information is useful in the remainder of the proof.

Now, fix i. The proof will proceed by induction. Define the induction hypothesis $P(n')$ for $n' \leq n$ as follows.

Definition 8.1. *$P(n')$: For any $\overline{\varepsilon}^{i} \in \Re^{n}$ and for any $m^{*} \in (0,1)^{n}$ with $\sum_{j} (m^{*})^{j} = 1$, there exists a unique $\widehat{\varepsilon} \in \Re^{n'}$ such that $(\widetilde{m}^{i}(\widehat{\varepsilon}, (\overline{\varepsilon}^{i}_{-i})^{-[n']})^{[n']} = (m^{*}_{-i})^{[n']}$.*

In other words, $P(n')$ says that for any value of the $\overline{\varepsilon}^{ij}$'s from $j = n' + 1$ to n and for any set of desired gross flows m^{*} from $j = 1$ to n', we can find exactly one choice of $\overline{\varepsilon}^{ij}$'s from $j = 1$ to n' (denoted $\hat{\varepsilon}$) that will provide exactly those desired gross flows. Where $P(n')$ holds, it will be useful to write the $\widehat{\varepsilon}$ as a function: $\widehat{\varepsilon}((\overline{\varepsilon}^{i}_{-i})^{-[n']}; (m^{*}_{-i})^{[n']})$.

Of course, the statement to be proved is simply $P(n)$. It is clear that $P(1)$ holds, since by (3.5) m_{i1} is continuous and strictly increasing in $\overline{\varepsilon}_{i1}$, $m_{i1} \to 0$ as $\overline{\varepsilon}_{i1} \to -\infty$ and $m_{i1} \to 1$ as $\overline{\varepsilon}_{i1} \to \infty$. Thus, the only task remaining is to show that $P(n')$ implies $P(n'+1)$.

Suppose that $P(n')$ holds, and so the $\widehat{\varepsilon}$ function defined above exists. Fix $(\overline{\varepsilon}^*_{-i})^{-[n']}$ and $(m^*_{-i})^{[n']}$. Consider the first n' elements of the $\tilde{m}^i$ function as a function of $(\overline{\varepsilon}^i)^{[n']}$, holding $(\overline{\varepsilon}^i)^{-[n']}$ constant. By (8.1), the derivatives of this function form an n'-square dominant diagonal matrix with positive elements on the main diagonal and negative elements off it. This implies that the inverse of that matrix exists and that it has all positive elements (see Theorems 1 and 4 in McKenzie (1960), respectively). This inverse is, then, the Jacobian of the $\widehat{\varepsilon}$ function with respect to $(m^*_{-i})^{[n']}$.

For any $\overline{\varepsilon}^{i,n'+1}$, define:

$$\mu(\overline{\varepsilon}^{i,n'+1}) \equiv (\widetilde{m}^i)(\widehat{\varepsilon}(\overline{\varepsilon}^{i,n'+1},(\overline{\varepsilon}^*_{-i})^{-[n'+1]};(m^*_{-i})^{[n']}),\overline{\varepsilon}^{i,n'+1},(\overline{\varepsilon}^*_{-i})^{-[n'+1]}),$$

the flow vector resulting from a given choice for $\overline{\varepsilon}^{i,n'+1}$, given that $\overline{\varepsilon}^{i,k}$ have been fixed for $k > n'+1$ and that $\overline{\varepsilon}^{i,k}$ for $k \leq n'$ are adjusted to keep the first n' elements of the flow vector equal to $(m^*_{-i})^{[n']}$. The μ function is differentiable by construction. The derivative of its first $n'+1$ elements is equal to:

$$\left(\frac{(\partial\tilde{m}^i)^{[n'+1]}}{(\partial\overline{\varepsilon}^i_{-i})^{[n'+1]}}\right)\begin{bmatrix}\frac{\partial\hat{\varepsilon}}{\partial\overline{\varepsilon}^{i,n'+1}}\\ 1\end{bmatrix} = \begin{bmatrix}\overrightarrow{0}\\ \frac{d\mu^{n'+1}}{d\overline{\varepsilon}^{i,n'+1}}\end{bmatrix}.$$

The left hand side of this equation is an $n'+1$-square matrix of derivatives multiplied by an $(n'+1)$-by-1 vector. The right hand side is an $(n'+1)$-by-1 vector that has n' zeroes, due to the definition of the $\hat{\varepsilon}$ function. Once again, by the properties of dominant diagonal matrices, the inverse of the first matrix on the left hand side exists and has only positive elements. Therefore, every element of the vector on the left-hand side has the same sign as $d\mu^{n'+1}/d\overline{\varepsilon}^{i,n'+1}$. Since $1 > 0$, this means that $d\mu^{n'+1}/d\overline{\varepsilon}^{i,n'+1} > 0$. Further, $d\hat{\varepsilon}/d\overline{\varepsilon}^{i,n'+1}$ is positive in each element.

From (3.5), we can see that $\mu^{n'+1} \to 0$ as $\overline{\varepsilon}^{i,n'+1} \to -\infty$. (For example, as $\overline{\varepsilon}^{i,n'+1} \to -\infty$, $F(\varepsilon^{n'+1} + \overline{\varepsilon}^{i,n'+1} - \overline{\varepsilon}^{i,n}) \to 0$ pointwise, so by the dominated convergence theorem $m^{i,n'+1} \to 0$.) Further, $\mu^k \to 0$ as $\overline{\varepsilon}^{i,n'+1} \to \infty$ for $k >$

$n' + 1$ (by a parallel argument), so $\mu^{n'+1} \to \left(1 - \sum_{j=1}^{n'} (m_{-i}^*)^j\right)$ as $\bar{\varepsilon}^{i,n'+1} \to \infty$. Therefore, by continuity, there exists a value of $\bar{\varepsilon}^{i,n'+1}$ such that

$$(\tilde{m}^i(\hat{\varepsilon}(\bar{\varepsilon}^{i,n'+1}, (\bar{\varepsilon}_{-i}^*)^{-[n'+1]}; (m_{-i}^*)^{[n']}), \bar{\varepsilon}^{i,n'+1}, (\bar{\varepsilon}_{-i}^*)^{-[n'+1]}))^{[n'+1]} = (m_{-i}^*)^{[n'+1]}.$$

Finally, since $d\mu^{n'+1}/d\bar{\varepsilon}^{i,n'+1} > 0$, as noted above, this value of $\bar{\varepsilon}^{i,n'+1}$ is unique. Thus, $P(n'+1)$ holds. ■

Proof of Proposition (3.4). Claim (i) is straightforward, since the planner could always set $D^{ii} \equiv 1$ for all i, which would ensure a non-negative value for (3.2) since $C^{ii} \equiv 0$. Claim (ii) follows from the continuity of the value function (trivially implied by its differentiability).

The proof of claim (iii) is as follows. Return to the original form of the problem, (3.2). Fix $L^* > \overline{L}$, and define $S^L(L^*) \equiv \{L \in \Re^n | L^i \geq 0; \sum_i L^i \in [0, L^*]\}$. If $L^{**} > L^*$, it is easy to see that for states in $S^L(L^*) \times S^S$, the value function that solves the Bellman equation with the state space limited to $S^L(L^*) \times S^S$ will agree with the function that solves it with the state space $S^L(L^{**}) \times S^S$. Thus, if we show that the Bellman equation derived for any finite L^* is concave in L, we are done. That will now be demonstrated.

For any $L \in S^L(L^*)$ and for any $n \times n$ matrix D of functions $D^{i,j} : \Re^n \mapsto [0,1]$, define

$$B(L, D) = \sum_{i=1}^{n} L_t^i \int \cdots \int \left(\sum_{j=1}^{n} D^{ij}(\varepsilon)(\varepsilon^j - C^{ij}) \right) \prod_{j=1}^{n} (f(\varepsilon^j) d\varepsilon^j).$$

This is the second term in the objective function. In addition, define the Bellman operator T on the space of bounded real functions on $S^L(L^*) \times S^s$ by:

$$T(W)(L, s) = \sup_D \{ \sum_{i=1}^{n} X^i(L, s) + B(L, D) + \beta E_{\tilde{s}}[W(\widetilde{L}, \tilde{s}) | s] \},$$

where $\widetilde{L}$ is determined from L and D by (3.6). A fixed point of T will be a solution to the Bellman equation, and by the usual logic of discounted dynamic programming, T is a contraction mapping, so that there is a unique fixed point, and it can be found as the limit of $T^k(W)$ as $k \to 0$ for any bounded function W.

Now consider a bounded and concave function W, and consider two different points in the state space, $a = (L_a, s)$ and $b = (L_b, s)$. In the optimization required in the definition of $T(W)$, denote the allocation rule chosen at state a by D_a, and the induced next-period labor allocation by $\widetilde{L}_a$, and similarly use D_b and $\widetilde{L}_b$ for state b. Now, consider the point $c = \alpha L_a + (1-\alpha)L_b$, for some $\alpha \in [0,1]$. Construct the allocation rule:

$$D_c^{ij}(\varepsilon) = [\alpha L_a^i D_a^{ij}(\varepsilon) + (1-\alpha)L_b^i D_b^{ij}(\varepsilon)]/L_c.$$

Since D_c is a weighted average of D_a and D_b within each cell, it satisfies (3.1) and is thus feasible. Note that:

$$\begin{aligned}
B(L_c, D_c) &= \sum_{i=1}^n L_c^i \int \cdots \int \left(\sum_{j=1}^n D_c^{ij}(\varepsilon)(\varepsilon^j - C^{ij}) \right) \prod_{j=1}^n (f(\varepsilon^j) d\varepsilon^j) \\
&= \sum_{i=1}^n \int \cdots \int \left(\sum_{j=1}^n (\alpha L_a^i D_a^{ij}(\varepsilon) + (1-\alpha)L_b^i D_b^{ij}(\varepsilon))(\varepsilon^j - C^{ij}) \right) \prod_{j=1}^n (f(\varepsilon^j) d\varepsilon^j) \\
&= \alpha B(L_a, D_a) + (1-\alpha) B(L_b, D_b).
\end{aligned}$$

Further, the next-period labor allocation vector that it induces is equal to $\alpha\widetilde{L}_a + (1-\alpha)\widetilde{L}_b$. We now have:

$$\begin{aligned}
T(W)(L_c, s) &\geq \sum_{i=1}^n X^i(L_c, s) + B(L_c, D_c) + \beta E_{\widetilde{s}}[W(\widetilde{L}_c, \widetilde{s})|s]\} \\
&= \sum_{i=1}^n X^i(L_c, s) + \alpha B(L_a, D_a) + (1-\alpha)B(L_b, D_b) + \beta E_{\widetilde{s}}[W(\widetilde{L}_c, \widetilde{s})|s]\} \\
&> \alpha T(W)(L_a, s) + (1-\alpha)T(W)(L_b, s).
\end{aligned}$$

The first inequality follows from optimization, and the fact that D_c is feasible. The last inequality follows from the concavity of X^i and W, and from the fact that D_a is optimal at point a and D_b is optimal at point b.

Therefore, if W is bounded and concave, so will be $T^k(W)$ for any k, and so must be the limit function, which is the true value function V. This completes the proof. ■

Proof of Proposition (3.5). Note that the derivative of U with respect to the choice variable is given by:

$$\frac{\partial U(L,s,\overline{\varepsilon})}{\partial \overline{\varepsilon}^{ii'}}$$
$$= -L^i \sum_{j\neq i'} \int (\varepsilon^j - C^{ij}) f(\varepsilon^j) f(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ii'}) \prod_{k\neq j,i'} F(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}) d\varepsilon^j$$
$$+ L^i \int_{-\infty}^{\infty} (\varepsilon^{i'} - C^{ii'}) f(\varepsilon^{i'}) \sum_{k\neq i'} f(\varepsilon^{i'} + \overline{\varepsilon}^{ii'} - \overline{\varepsilon}^{ik}) \prod_{l\neq i',k} F(\varepsilon^{i'} + \overline{\varepsilon}^{ii'} - \overline{\varepsilon}^{il}) d\varepsilon^{i'}.$$

Using the change of variables $\varepsilon = \varepsilon^j - \overline{\varepsilon}^{ii'} + \overline{\varepsilon}^{ij}$ on the first integral and rearranging yields:

$$\begin{aligned} \frac{\partial U(L,s,\overline{\varepsilon})}{\partial \overline{\varepsilon}^{ii'}} &= L^i \sum_{j\neq i'} (\overline{\varepsilon}^{ii'} - \overline{\varepsilon}^{ij} + C^{ii'} - C^{ij}) \frac{\partial m^{ij}}{\partial \overline{\varepsilon}^{ii'}} \\ &= L^i \sum_{j=1}^{n} (-\overline{\varepsilon}^{ij} - C^{ij}) \frac{\partial m^{ij}}{\partial \overline{\varepsilon}^{ii'}}. \end{aligned}$$

(The equality follows, first, because the term in parentheses equals zero when $j = i'$, so we can lift the restriction that $j \neq i'$ without affecting the equation; and second, the sum of derivatives of the flows across all cells resulting from a change in $\overline{\varepsilon}^{ii'}$ must equal zero.) The first order condition for the Bellman equation is, then:

$$L^i \sum_{j=1}^{n} \left(-\overline{\varepsilon}^{ij} - C^{ij} + \beta E \frac{\partial \widetilde{V}}{\partial \widetilde{L}^j} \right) \frac{\partial m^{ij}}{\partial \overline{\varepsilon}^{ii'}} = 0.$$

Define the function $\widetilde{\varepsilon}^i$ as the inverse of the function $\widetilde{m}^i$ defined in the beginning of the proof of Proposition (3.3). Then the first order condition implies, if $i \neq 1$:

$$\begin{aligned} & \sum_{i'\neq i} \left(L^i \sum_{j=1}^{n} \left(-\overline{\varepsilon}^{ij} - C^{ij} + \beta E \frac{\partial \widetilde{V}}{\partial \widetilde{L}^j} \right) \frac{\partial m^{ij}}{\partial \overline{\varepsilon}^{ii'}} \frac{\partial \widetilde{\varepsilon}^{ii'}}{\partial m^{i1}} \right) \\ = \; & L^i \sum_{j=1}^{n} \left(-\overline{\varepsilon}^{ij} - C^{ij} + \beta E \frac{\partial \widetilde{V}}{\partial \widetilde{L}^j} \right) \sum_{i'\neq i} \frac{\partial m^{ij}}{\partial \overline{\varepsilon}^{ii'}} \frac{\partial \widetilde{\varepsilon}^{ii'}}{\partial m^{i1}} = 0 \end{aligned}$$

Now, note that

$$\sum_{i' \neq i} \frac{\partial m^{ij}}{\partial \overline{\varepsilon}^{ii'}} \frac{\partial \widetilde{\varepsilon}^{ii'}}{\partial m^{i1}}$$

takes a value of 1 if j equals 1, -1 if j equals i, and zero otherwise. Thus, the first order condition reduces to:

$$L^1 \left(-\overline{\varepsilon}^{i1} - C^{i1} + \beta E \frac{\partial \widetilde{V}}{\partial \widetilde{L}^1} + \overline{\varepsilon}^{ii} + C^{ii} - \beta E \frac{\partial \widetilde{V}}{\partial \widetilde{L}^i} \right) = 0, \text{ or}$$

$$\overline{\varepsilon}^{i1} + C^{i1} = \beta E \left(\frac{\partial \widetilde{V}}{\partial \widetilde{L}^1} - \frac{\partial \widetilde{V}}{\partial \widetilde{L}^i} \right).$$

This equation says that the marginal cost of moving a worker from i to 1 is equal at the optimum to the expected discounted marginal benefit of doing so. This can be repeated for any pair of cells i and j with $i \neq j$, to yield the indicated condition. ■

Proof of Proposition (3.6). Using (3.8) and (3.4), we have:

$$\begin{aligned} &\frac{\partial V(L,s)}{\partial L^i} \\ = \quad & X_1^i + \sum_{j=1}^{n} \left(\int_{-\infty}^{\infty} (\varepsilon^j - C^{ij}) f(\varepsilon^j) \prod_{k \neq j} F(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}) d\varepsilon^j \right) + \beta \sum_{j=1}^{n} m^{ij} \frac{\partial \tilde{V}}{\partial \widetilde{L}^j}, \end{aligned}$$

where $\widetilde{V}$ stands for $E[V(\widetilde{L}, \widetilde{s})|s]$ from (3.8). Rearranging, this becomes

$$\begin{aligned} & X_1^i + \sum_{j=1}^{n} \left(\int_{-\infty}^{\infty} \varepsilon^j f(\varepsilon^j) \prod_{k \neq j} F(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}) d\varepsilon^j \right) \\ & + \sum_{j=1}^{n} m^{ij} \left(-C^{ij} + \beta \left(\frac{\partial \tilde{V}}{\partial \widetilde{L}^j} - \frac{\partial \tilde{V}}{\partial \widetilde{L}^i} \right) \right) + \beta \frac{\partial \tilde{V}}{\partial \widetilde{L}^i}, \end{aligned}$$

which from (3.9) and the definition of the gross flows, (3.5), becomes

$$X_1^i + \sum_{j=1}^{n} \left(\int_{-\infty}^{\infty} \left(\varepsilon^j + \overline{\varepsilon}^{ij} \right) f(\varepsilon^j) \prod_{k \neq j} F(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}) d\varepsilon^j \right) + \beta \frac{\partial \tilde{V}}{\partial \widetilde{L}^i}.$$

This is the indicated condition. ■

Proof of Proposition (5.1). Suppose that W^* and W^{**} are two bounded, concave value functions with $dW^*(\overline{L}-L^2,L^2)/dL^2 \geq dW^{**}(\overline{L}-L^2,L^2)/dL^2$ for all $L^2 \in (0,\overline{L}]$ (call this the 'weak derivative property'), and let T be the operator on value functions defined by the planner's Bellman equation. Then we claim that $T(W^*)$ and $T(W^{**})$ are both bounded and continuous with $dT(W^*)(\overline{L}-L^2,L^2)/dL^2 \geq dT(W^{**})(\overline{L}-L^2,L^2)/dL^2$ for all $L^2 \in (0,\overline{L}]$.

The boundedness of $T(W^*)$ and $T(W^{**})$ is immediate, and their concavity can be proven with the same argument as was used in the proof of Proposition (3.4). For the derivative property, note that the first-order condition for solving the Bellman equation with the function W^{**} is:

$$\overline{\varepsilon}^{12} + C^{12} = \beta[dW^{**}(\overline{L}-\widetilde{L}^2,\widetilde{L}^2)/d\widetilde{L}^2],$$

where a tilde indicates next-period variables computed from the gross flow matrix. Given that $\overline{\varepsilon}^{21}_{t+1} = -\overline{\varepsilon}^{12}_{t+1} - C^{12} - C^{21}$ at all times (see (3.9)), we can think of $\overline{\varepsilon}^{21}_{t+1}$ as a decreasing function of $\overline{\varepsilon}^{12}_{t+1}$. Thus, an increase in $\overline{\varepsilon}^{12}$ will increase m^{12} and decrease m^{21}, increasing the next-period value of L^2. By the concavity of W^{**}, this will decrease the value of the right-hand side of the first-order condition. Thus, the right-hand side of the condition is a downward-sloping curve in $\overline{\varepsilon}^{12}$. At the same time, the left hand side of the condition is an upward-sloping line in $\overline{\varepsilon}^{12}$. As a result, for a given value of L^2, anything that shifts the right-hand side of the first-order condition down will result in a lower value of $\overline{\varepsilon}^{12}$. Therefore, for a given value of L^2, the solution to the first-order condition with W^{**} will yield a lower value of $\overline{\varepsilon}^{12}$, and thus a higher value of $\overline{\varepsilon}^{21}$, along with a lower value for the right-hand side of the first-order condition, than will the solution with W^*. But then applying the envelope condition (3.10) to $T(W^*)$ and $T(W^{**})$, it is clear that the weak derivative property holds for $T(W^*)$ and $T(W^{**})$. This proves the claim.

Therefore, from any initial bounded and concave W^* and W^{**} satisfying the weak derivative property, $T^k(W^*)$ and $T^k(W^{**})$ will also be bounded and concave and satisfy the derivative property for any k, and so the property holds in the limit as $k \to \infty$. Thus, the value functions V^* and V^{**} also satisfy the weak derivative property.

From here there is one step required to show that the value functions satisfy the *strong* derivative property. Considering the first-order conditions again, this time for V^* and V^{**} respectively, the curve-shifting logic used in the proof of the claim above shows that for a given value of L^2, the value of $\overline{\varepsilon}^{12}$ chosen with V^*

will be at least as great as that chosen with V^{**}. Therefore, again looking at the envelope condition (3.10) and noting that $X_1^{2*}(L^2) > X_1^{2**}(L^2)$ for all $L^2 > 0$, the strong derivative condition is immediate. ■

Proof of Proposition (5.2). Fix $\delta > 0$. Rewrite the planner's objective function (3.4):

$$X(L_t, s_t) + \sum_{ij} L_t^i \int \varepsilon^j \prod_{k \neq j} G_\delta(\overline{\varepsilon}_t^{ij} - \overline{\varepsilon}_t^{ik} + \varepsilon_t^j) g_\delta(\varepsilon^j) d\varepsilon^j - \sum_{i,j} L_t^i m_\delta^{ij}(\overline{\varepsilon}_t) C^{ij},$$

where m_δ^{ij} denotes the gross flow from i to j as calculated from (3.5) using the distribution G_δ, and, as before $\overline{\varepsilon}^{ii} = 0 \forall i$. We can rewrite this function once again as follows.

$$U_\delta(L, s, \widetilde{\varepsilon}) \equiv$$
$$X(L, s) + \sum_{i,j} L^i \int \varepsilon^j \prod_{k \neq j} G_\delta(\delta(\widetilde{\varepsilon}^{ij} - \widetilde{\varepsilon}^{ik}) + \varepsilon^j) g_\delta(\varepsilon^j) d\varepsilon^j - \sum_{i,j} L^i m_\delta^{ij}(\delta\widetilde{\varepsilon}) C^{ij},$$

where $\widetilde{\varepsilon}$ is an n-square matrix of real numbers with $\widetilde{\varepsilon}^{ii} = 0$. In other words, $\widetilde{\varepsilon}$ is simply $\overline{\varepsilon}$, scaled down by a factor of δ.

Since

$$\begin{aligned} m_\delta^{ij}(\delta\widetilde{\varepsilon}) &= \int \prod_{k \neq j} G_\delta(\delta(\widetilde{\varepsilon}^{ij} - \widetilde{\varepsilon}^{ik}) + \varepsilon^j) g_\delta(\varepsilon^j) d\varepsilon^j \\ &= \int \prod_{k \neq j} G_1(\widetilde{\varepsilon}^{ij} - \widetilde{\varepsilon}^{ik} + \frac{\varepsilon^j}{\delta}) g_1(\frac{\varepsilon^j}{\delta})(\frac{1}{\delta}) d\varepsilon^j \\ &= \int \prod_{k \neq j} G_1(\widetilde{\varepsilon}^{ij} - \widetilde{\varepsilon}^{ik} + \varepsilon) g_1(\varepsilon) d\varepsilon \\ &= m_1^{ij}(\widetilde{\varepsilon}), \end{aligned}$$

the gross flows resulting from any given choice of $\widetilde{\varepsilon}$ are independent of δ.

Further,

$$\begin{aligned}
&\sum_{i,j} L^i \int \varepsilon^j \prod_{k\neq j} G_\delta(\delta(\widetilde{\varepsilon}^{ij} - \widetilde{\varepsilon}^{ik}) + \varepsilon^j) g_\delta(\varepsilon^j) d\varepsilon^j \\
= \;& \delta \sum_{i,j} L^i \int \frac{\varepsilon^j}{\delta} \prod_{k\neq j} G_1(\widetilde{\varepsilon}^{ij} - \widetilde{\varepsilon}^{ik} + \frac{\varepsilon^j}{\delta}) g_1(\frac{\varepsilon^j}{\delta})(\frac{1}{\delta}) d\varepsilon^j \\
= \;& \delta \sum_i L^i A^i(\widetilde{\varepsilon}^i),
\end{aligned}$$

where

$$A^i(\widetilde{\varepsilon}^i) \equiv \sum_j \int \varepsilon \prod_{k\neq j} G_1(\widetilde{\varepsilon}^{ij} - \widetilde{\varepsilon}^{ik} + \varepsilon) g_1(\varepsilon) d\varepsilon.$$

Each of these A^i functions takes a unique maximum at $\widetilde{\varepsilon}^i = 0$. To see this, consider a sample of n independent draws from the distribution G_1, and call the realized values $\varepsilon^1, \ldots \varepsilon^n$. The function $A^i(\widetilde{\varepsilon}^i)$ is the expectation of the j^*th of these, where j^* is the value of j that maximizes $\{\widetilde{\varepsilon}^{ij} + \varepsilon^j\}$. On the other hand, $A^i(0)$ is simply the expectation of the highest of the ε^j's. Thus, $A^i(0)$ must be higher.

We can now rewrite the objective function once again:

$$U_\delta(L, s, \widetilde{\varepsilon})/\delta = \sum_i L^i A^i(\widetilde{\varepsilon}^i) + [X(L, s) - \sum_{i,j} L^i m_1^{ij}(\widetilde{\varepsilon}) C^{ij}]/\delta. \qquad (8.2)$$

The maximization of (3.3) is, of course, equivalent to maximizing the expected present discounted value of $U_\delta(L, s, \widetilde{\varepsilon})/\delta$. Further, we can speak in terms of the optimal choice of $\widetilde{\varepsilon}$ in each state instead of the optimal choice of $\overline{\varepsilon}$ without making any substantive difference.

Fix $\Delta > 0$. Let $\widehat{\Delta} = \sum_i L^i A^i(0) - \sup_{|\widetilde{\varepsilon}| \geq \Delta} \sum_i L^i A^i(\widetilde{\varepsilon}) > 0$, where $|\widetilde{\varepsilon}|$ indicates the absolute value of the element of $\widetilde{\varepsilon}$ that is farthest from zero. (Think of $\widehat{\Delta}$ as the minimum loss from having $\widetilde{\varepsilon}$ a distance Δ away from its optimum of 0.) The point will be to demonstrate that if δ is large enough, we will have $|\widetilde{\varepsilon}| < \Delta$, regardless of the value of L and s.

From (2.1) and the fact that $\sum_i L^i \equiv \overline{L}$, the last two terms of (8.2) can be made uniformly arbitrarily small by choosing δ sufficiently high. Choose δ high enough that those two terms are always less than $(1 - \beta)\widehat{\Delta}/2$ in absolute value.

Now, suppose that the optimal rule for choosing $\widetilde{\varepsilon}$ has at some state (L^*, s^*) a value of $\widetilde{\varepsilon}$ with $|\widetilde{\varepsilon}| > \Delta$. Now, replace that rule with one that is identical except that at that state, and at all other states after that state has once been reached, $\widetilde{\varepsilon}$ is set equal to 0. In the first period in which the change takes effect, that would increase the value of the first term of (8.2) by at least $\widehat{\Delta}$. Thereafter, it could not reduce the value of that term, because with $\widetilde{\varepsilon} = 0$, that term would be at its maximum. On the other hand, in the first period of the change or in any subsequent period, the second two terms together could fall by at most $(1-\beta)\widehat{\Delta}/2$, so the expected present discounted value of the reduction in those terms would be at most $[(1-\beta)\widehat{\Delta}/2]/(1-\beta) = \widehat{\Delta}/2$. Thus, the change in the value of the objective function due to the change in rule evaluated at the state (L^*, s^*) would be at least equal to $\widehat{\Delta} - \widehat{\Delta}/2 = \widehat{\Delta}/2 > 0$. This contradicts the assumption that the initial rule was optimal.

Thus, we have that $\widetilde{\varepsilon}$ as a function of L and s converges uniformly to the constant 0 as $\delta \to \infty$. Since the function m_1 is continuous and

$$\begin{aligned} m_1^{ij}(0) &= \int \prod_{k \neq j} G_1(\varepsilon) g_1(\varepsilon) d\varepsilon \\ &= \frac{1}{n} G_1(\varepsilon)^n |_{-\infty}^{\infty} \\ &= \frac{1}{n}, \end{aligned}$$

we conclude that $m_1^{ij}(\widetilde{\varepsilon}(L,s))$ converges to the constant $1/n$ uniformly as $\delta \to \infty$. ■

References

[1] Borjas, George J. (1995). "The Economics of Immigration." *Journal of Economic Literature* 32, pp. 1667-717.

[2] Dehejia, Vivek (1997). "Will Gradualism work when Shock Therapy Doesn't?" CEPR Discussion Paper No. 1552, Centre for Economic Policy Research, London, February.

[3] Dixit, Avinash (1993). "Prices of Goods and Factors in a Dynamic Stochastic Economy." In Wilfred J. Ethier, Elhanan Helpman and J. Peter Neary (eds.), *Theory, Policy and Dynamics in International Trade: Essays in Honor of Ronald W. Jones.* Cambridge, U.K.: Cambridge University Press.

[4] Dixit, Avinash and Rafael Rob (1994). "Switching Costs and Sectoral Adjustments in General Equilibrium with Uninsured Risk." *Journal of Economic Theory* 62, pp. 48-69.

[5] Freund, Caroline, and John McLaren (1999). "On the Dynamics of Trade Diversion: Evidence from Four Trade Blocks". *International Finance Discussion Paper* 637, Board of Governors of the Federal Reserve System, Washington, D.C.

[6] Gale, David and Hukukane Kikaido (1965). "The Jacobian Matrix and Global Univalence of Mappings." *Mathematische Annalen* 159, pp.81-93.

[7] Hanson, Gordon H. and Matthew J. Slaughter (1999). "The Rybczynski Theorem, Factor-Price Equalization, and Immigration: Evidence from U.S. States." *NBER Working Paper* 7074 (April).

[8] Irwin, Douglas A. (1996). "Industry or Class Cleavages over Trade Policy? Evidence from the British General Election of 1923." In Feenstra, Robert C., Gene M. Grossman, and Douglas A. Irwin (eds.), *The Political Economy of Trade Policy: Papers in Honor of Jagdish Bhagwati*. Cambridge, MA: MIT Press. pp. 53-75.

[9] Jovanovic, Boyan and Robert Moffitt (1991). "An Estimate of a Sectoral Model of Labor Mobility," *Journal of Political Economy*, 98:4, pp. 827-852.

[10] Magee, Stephen P. (1989). "Three Simple Tests of the Stolper-Samuelson Theorem." Ch. 7 in Stephen P. Magee, William A. Brock, and Leslie Young, *Black Hole Tariffs and Endogenous Policy Theory*, New York: Cambridge University Press, 1989.

[11] Matsuyama, Kiminori (1992). "A Simple Model of Sectoral Adjustment." *Review of Economic Studies,* 59, pp. 375-88.

[12] McKenzie, Lionel (1960). "Matrices with Dominant Diagonals and Economic Theory," Chapter 4 in Kenneth J. Arrow, Samuel Karlin and Patrick Suppes, *Mathematical Methods in the Social Sciences*, 1959, pp. 47-62. Stanford, California: Stanford University Press.

[13] McLaren, John (1997). "Size, Sunk Costs, and Judge Bowker's Objection to Free Trade," *American Economic Review,* 87:3 (June), pp. 400-20.

[14] _____(2002) "A Theory of Insidious Regionalism." Forthcoming, *Quarterly Journal of Economics.*

[15] Mussa, Michael (1978). "Dynamic Adjustment in the Heckscher-Ohlin-Samuelson Model," *Journal of Political Economy,* 86:5, pp. 775-91.

[16] _____(1982). "Government Policy and the Adjustment Process," in Jagdish N. Bhagwati (ed.), *Import Competition and Response*, Chicago: University of Chicago Press.

[17] Neary, Peter (1985). "Theory and Policy of Adjustment in an Open Economy," in David Greenaway, *Current Issues in International Trade: Theory and Policy*, Houndmills, UK: Macmillan, pp.43-61.

[18] Patel, Jagdish, C. H. Kapadia, and D. B. Owen (1976). *Handbook of Statistical Distributions*, New York: Marcel Dekker, Inc.

[19] Rappaport, Jordan (2000). "Why are Population Flows so Persistent?" Federal Reserve Bank of Kansas City Working Paper RWP 99-13.

[20] Scheve, Kenneth F. and Matthew J. Slaughter (1998). "What Explains Individual Trade-Policy Preferences?" NBER Working paper #6531, (April).

[21] _____(1999). "Labor-Market Competition and Individual Preferences Over Immigration Policy." NBER Working Paper #6946 (February).

[22] Sjaastad, Larry A. (1962) "The Costs and Returns of Human Migration." *Journal of Political Economy,* 70: 5, pt. 2, pp. S80-S93.

[23] Slaughter, Matthew J. (1998). "International Trade and Labor-Market Outcomes," *Economic Journal*, 108:450 (September), pp. 1452-1462.

[24] Staiger, Robert and Guido Tabellini (1999). "Do GATT Rules Help Governments Make Domestic Commitments?" *Economics and Politics*,11:2 (July), pp.109-144.

[25] Stokey, Nancy L. and Robert E. Lucas, Jr. (1989). *Recursive Methods in Economic Dynamics.* Cambridge, MA: Harvard University Press.

[26] Topel, Robert H. (1986) "Local Labor Markets." *Journal of Political Economy,* 94: 3, pt. 2, pp. S111-S143.