WPS6480

Policy Research Working Paper 6480

PPML Estimation of Dynamic Discrete Choice Models with Aggregate Shocks

Erhan Artuc

The World Bank
Development Research Group
Trade and International Integration Team
June 2013

Policy Research Working Paper 6480

Abstract

This paper introduces a computationally efficient method for estimating structural parameters of dynamic discrete choice models with large choice sets. The method is based on Poisson pseudo maximum likelihood (PPML) regression, which is widely used in the international trade and migration literature to estimate the gravity equation. Unlike most of the existing methods in the literature, it does not require strong parametric assumptions on agents' expectations, thus it can accommodate macroeconomic and policy shocks. The regression requires count data as opposed to choice probabilities; therefore it can handle sparse decision transition matrices caused by small sample sizes. As an example application, the paper estimates sectoral worker mobility in the United States.

This paper is a product of the Trade and Integration Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at eartuc@worldbank.org.

PPML Estimation of Dynamic Discrete Choice Models with Aggregate Shocks

Erhan Artuc*

Keywords: Poisson Pseudo Maximum Likelihood, Labor Mobility, Migration, Discrete Choice Models, Gravity Equation.

JEL Codes: C25, F22, J61, J62, F16.

*The views in this paper are the author's and not those of the World Bank Group or any other institution. Artuc: World Bank, Trade and International Integration Unit, Development Economics Research Group (Economic Policy), 1818 H Street, NW Washington DC, 20433 USA; eartuc@worldbank.org. I thank Jim Anderson, Chad Bown, Irene Brambilla, David Kaplan, Hiau Looi Kee, John Kennan, John McLaren, Caglar Ozden, Daniel Lederman, Guido Porto, Ray Robertson, Diego Rojas and Yoto Yotov for their comments. All errors and omissions are mine.

Poisson pseudo maximum likelihood regression (henceforth PPML) within the context of the gravity equation, has become very popular in international trade and migration literature. It was introduced by Gourieroux, Monfort and Trognon (1984). Then, more recently, Santos Silva and Tenreyro (2006) showed that it is a simple but powerful method for estimating bilateral resistance parameters of the gravity equation. After these two seminal papers, it has become one of the standard tools in the international economics literature, widely used to explain trade, and more recently migration flows[1]. This paper extends this popular method further and shows how it can be used to estimate structural dynamic discrete choice models by adding a linear reduced form regression step. This novel method can handle models with large choice sets, heterogeneity, and aggregate shocks. Our approach is an intuitive combination of well known and widely used methods, therefore it imposes little set-up cost to the econometrician and can utilize standard statistical software.

The method has two steps: First, we run PPML regression using discrete choice data, similar to the gravity equation estimation, to estimate expected values that appear in the Bellman equations. Second, we construct a linear regression equation by plugging the estimated expected values into the Bellman equation that characterizes the dynamic decision making process of agents. In the second step, we estimate distributional and utility flow parameters of the discrete choice model. Since we estimate expected values rather than calculating value functions by iteration or backward solution, expectations of agents are fully accounted for even when they are not quantifiable by the econometrician. Both regressions are based on orthogonality conditions, rather than maximum likelihood, therefore the distributions of payoff streams or aggregate shocks are not required for the estimation. The estimated system does not need to be at the steady state. In fact, the steady state may not even exist in the presence of macroeconomic and policy shocks. The orthogonality conditions we use have analytical derivatives, therefore the method we use is much faster than maximum

[1]Two recent examples from the migration literature that use gravity equation are Beine, Docquier and Ozden (2011) and Grogger and Hanson (2011). In the trade literature, Olivero and Yotov (2011) generalize gravity equation in a dynamic framework but estimate trade flows at the steady state without considering a discrete choice specification.

likelihood based methods and it allows us to estimate a large number of parameters.

Accounting for aggregate shocks is an important challenge for the estimation of dynamic discrete choice models. In the literature, the most common methods are based on maximum likelihood estimation (henceforth ML) using backwards solution or conditional choice probabilities. ML estimation requires strong distributional assumptions on aggregate shocks, thus on workers' expectations about payoffs. In the literature, the most common assumption is the absence of aggregate shocks; because it is very difficult to rigorously model transmission of aggregate shocks into the payoff streams and workers' expectations[2]. In contrast to ML estimation, our method does not require distributional assumptions on aggregate shocks or workers' expectations, except rationality. Therefore, using this novel method, international and internal migration, sectoral labor mobility, occupational mobility, and other dynamic discrete choice models can be estimated efficiently in the presence of macroeconomic and policy shocks.

The two most recent papers that address similar discrete choice problems are Anderson (2011) and Artuc, Chaudhuri and McLaren (2010). Anderson (2011) shows how the gravity equation can be considered as an equilibrium condition for discrete choice problems, and it can be estimated with PPML regression. After the estimation step, he solves the structural push and pull parameters from the PPML regression coefficients along with multilateral resistance parameters. Technically, our first step regression is similar to a gravity equation, as in Anderson (2011). Different from him, we interpret PPML regression fixed effects as expected values in the Bellman equation that gives the optimality condition for the underlying discrete choice model. We estimate parameters of the Bellman equation rather than the gravitational push and pull parameters.

[2]For example, after the recent housing crisis of 2007, the demand for unskilled labor in the construction sector decreased significantly. This negative shock potentially affected expected payoffs of workers in the construction sector and welfare of low-skill immigrants in general. Consequently, the construction sector shrank significantly and some states such as Alabama and Arizona passed strict anti-immigration laws. Any migration or sectoral mobility model that is estimated via ML has to contain distributional assumptions regarding such policy and macroeconomic shocks that are difficult to quantify.

Artuc, Chaudhuri and McLaren (2010) derive an equilibrium condition for workers' sectoral choice, that is in essence an Euler equation. Similar to ours, their method also allows aggregate shocks. They impute workers' values from observed gross flows. However, their expected value imputation method does not allow sparsity in the transition matrices of flows from one decision to another. This limits the number of choices, and also creates problems in incorporating heterogeneity. Furthermore, it increases the standard errors, preventing estimation of detailed versions of their model. Our method is free from these limitations.

The idea of imputing values from conditional choice probabilities (henceforth CCP) was first introduced by Hotz and Miller (1993), a milestone for discrete choice models due to their inversion theorem. They consider stationary problems and estimate choice probabilities using a non-parametric method. Aguirregabiria and Mira (2002) introduced a new algorithm, "Nested Pseudo Maximum Likelihood" (henceforth NPM) that combines a CCP with an iterative step to improve efficiency of CCP in small samples[3]. More recently, in their seminal paper, Arcidiacono and Miller (2011) combine CCP with the expectation-maximization (henceforth EM) algorithm to allow certain non-stationary processes and unobserved heterogeneity. Similar to Arcidiacono and Miller (2011), our estimation procedure can use an EM loop to account for unobserved heterogeneity. In this paper we argue that PPML is a convenient and efficient alternative to CCP for problems with large choice sets or a small sample. However, it is not a substitute for EM or NPM, and it can be used in combination with EM and NPM instead of non-parametric CCP estimation.

Our procedure has two major differences with the other non-iterative dynamic programming methods in the literature: First, we use Poisson regression rather than the Hotz and Miller (1993) inversion equation to estimate expected values. Second, our method does not rely on maximum likelihood estimation but orthogonality conditions. Therefore, we do not need distributional assumptions for aggregate shocks. Thanks to the linearity of the estimating equations, it is well suited for problems with a large number of choices and structural

[3]It is essentially an intuitive combination of Hotz and Miller (1993) and Rust (1987). See Aguirregabiria and Mira (2010) for an extensive survey of the literature.

parameters. Our method is computationally efficient and can utilize standard statistical software, widely used to estimate gravity equations in the context of migration and trade flows[4]. Similar to the other non-iterative solution methods, the state space has to be small compared to the backwards solution methods.

In the next section, we present a representative discrete choice model that can be estimated with our method. In the following sections, we summarize our estimation strategy, and provide an example application, and present simulation results.

1 Model

Consider an economy with infinitely-lived L agents and N sectors, where each agent is in a discrete state $s \in S$. Sectors can be industries, occupations, cities, countries, or any combination of such choices, while the state could be the type of agent such as education level, gender, age or other individual characteristics. It is also possible to consider economic policies as a part of state space, such as trade policy, migration policy, or education policy. We can also incorporate unobserved types, which is omitted from this section for the sake of clarity.

A type s agent chooses a sector $i \in \{1, 2, 3, .., N\}$ in the end of period $t-1$, and receives instantaneous utility $u_t^{i,s}$ at time t defined as

$$u_t^{i,s} = w_t^{i,s} + \eta^{i,s}, \tag{1}$$

where $w_t^{i,s}$ is the observed sector specific random payoff common to all type s agents working in sector i with finite moments, and $\eta^{i,s}$ is the unobserved sector specific iid utility shock also common to all type s agents. Hence, the state of each agent can be summarized with the pair (i, s) where s is the type and i is the current sector.

[4]PPML estimation is based on an orthogonality condition, which has an analytical derivative. This computational convenience makes the estimation process much faster than alternatives. For example, it converges within minutes even with hundreds of choices and many structural parameters.

We assume that only $w_t^{i,s}$ is observed by the econometrician, $\eta^{i,s}$ is known by agents, but not by the econometrician. All agents are risk neutral, have rational expectations and a common discount factor $\beta < 1$. The expected future payoff streams can change over time, $E_{t+1} w_{t+n}^{i,s} \neq E_t w_{t+n}^{i,s}$ for $n \geq 1$. The present discounted choice-specific utility of agent l is equal to

$$U_t^{i,s,l} = w_t^{i,s} + \eta^{i,s} + \max_j \left\{ \beta E_t V_{t+1}^{j,s'} - C_t^{ij,s} - \varepsilon_t^{j,l} \right\}, \tag{2}$$

where $C_t^{ij,s} + \varepsilon_t^{j,l}$ is the cost of choosing sector j, for type s agent l who is currently in sector i. The "moving cost" has two components, a deterministic part, $C_t^{ij,s}$, common to all type s agents, and a random part, $\varepsilon_t^{j,l}$, specific to agent l. All type s agents are identical except for their individual moving cost shock $\varepsilon_t^{j,l}$. We assume that $C_t^{ii,s} = 0$, which means the fixed component of moving cost is zero for stayers.

The timing of events is as follows: 1. Agents learn values of $w_t^{i,s}$ once they receive it. 2. Then, in the end of time t, they learn the random component of "moving cost," $\varepsilon_t^{j,l}$, for every $j = 1, .., N$, and choose the next period sector (based on expected stream of future payoffs and moving costs). 3. Agents pay the moving cost, $C_t^{ij,s} + \varepsilon_t^{j,l}$, where j is the chosen sector. 4. Period $t + 1$ starts, and the cycle repeats itself.

After taking expectation of (2) with respect to agent specific shocks, the choice specific value function can be expressed as

$$V_t^{i,s,l} = w_t^{i,s} + \eta^{i,s} + E_t \max_j \left\{ \beta \sum_{s' \in S} \pi(s, s') V_{t+1}^{j,s',l} - C_t^{ij,s} - \varepsilon_t^{j,l} \right\}, \tag{3}$$

where $\pi(s, s')$ is the probability of switching from type s to type s'. We assume that $\pi(s, s')$ is exogenous[5]. Henceforth, we drop the agent superscript l for notational convenience.

We can rearrange the value function as

$$V_t^{i,s} = w_t^{i,s} + \eta^{i,s} + \beta \tilde{V}_{t+1}^{i,s} + E_t \max_j \{ \varepsilon_t^j + \bar{\varepsilon}_t^{ij,s} \},$$

[5]It is possible to endogenize this transition matrix, but is out of scope of this paper.

where

$$\overline{\varepsilon}_t^{ij,s} = [\beta\widetilde{V}_{t+1}^{j,s} - \beta\widetilde{V}_{t+1}^{i,s}] - C_t^{ij,s},$$

and

$$\widetilde{V}_{t+1}^{i,s} = \sum_{s' \in S} \pi\left(s,s'\right) E_t V_{t+1}^{i,s'}. \tag{4}$$

Then, the choice specific values can be written as

$$V_t^{i,s} = w_t^{i,s} + \eta^{i,s} + \beta\widetilde{V}_{t+1}^{i,s} + \Omega_t^{i,s}. \tag{5}$$

The option value $\Omega_t^{i,s}$ is equal to

$$\Omega_t^{i,s} = \sum_{j=1}^{N} \int_{-\infty}^{\infty} (\varepsilon^j + \overline{\varepsilon}_t^{ij,s}) f(\varepsilon^j) \prod_{k \neq j} F(\varepsilon^j + \overline{\varepsilon}_t^{ij,s} - \overline{\varepsilon}_t^{ik,s}) d\varepsilon^j,$$

where $F(\varepsilon)$ is the cumulative distribution function and $f(\varepsilon)$ is the probability density function of the moving cost shocks. The option value, Ω_t^i, is the extra utility generated by being able to change sectors. As moving cost C_t^{ij} increases, the option value decreases, and it diminishes to zero when the moving cost goes to infinity. The option value function is crucial for the implementation of estimation process since it can be solved analytically under certain distributional assumptions.

Assume that ε_t^i is distributed iid extreme value type I with location parameter $-\nu\gamma$, scale parameter ν, and cdf $F\left(\varepsilon\right) = \exp\left(-\exp\left(-\varepsilon/\nu - \gamma\right)\right)$, where $E\left(\varepsilon\right) = 0$, $Var\left(\varepsilon\right) = \pi^2\nu^2/6$ and γ is the Euler's constant.

Assume that $m_t^{ij,s}$ is equal to the ratio of type s agents who switch from sector i to sector j. This can be interpreted as gross flows from i to j, or the probability of choosing j conditional on (i,s). The total number of agents moving from i to j is equal to $y_t^{ij,s} = L_t^{i,s} m_t^{ij,s}$, where $L_t^{i,s}$ is the number of type s agents who are in i at time t. $y_t^{ij,s}$ can be interpreted as number of people migrating from one city to another, changing occupation, changing industry, etc.

Thanks to the extreme value distribution and McFadden (1973), the gross flow $m_t^{ij,s}$ is equal to

$$m_t^{ij,s} = \frac{\exp\left(\left(\beta\widetilde{V}_{t+1}^{j,s} - \beta\widetilde{V}_{t+1}^{i,s} - C_t^{ij,s}\right)\frac{1}{\nu}\right)}{\sum_{k=1}^{N}\exp\left(\left(\beta\widetilde{V}_{t+1}^{k,s} - \beta\widetilde{V}_{t+1}^{i,s} - C_t^{ik,s}\right)\frac{1}{\nu}\right)}, \tag{6}$$

and we can show that the option value[6] is equal to

$$\Omega_t^{i,s} = \nu \log \sum_{k=1}^{N}\exp\left(\left(\beta\widetilde{V}_{t+1}^{k,s} - \beta\widetilde{V}_{t+1}^{i,s} - C_t^{ik,s}\right)\frac{1}{\nu}\right). \tag{7}$$

Note that we could use an expression similar to (7) to construct a CCP representation of the Bellman equation, because $\Omega_t^{i,s} = -\nu \log m_t^{ii}$ and m_t^{ij} is a conditional choice probability, see Appendix C for the details. We do not use the method CCP or the Hotz-Miller inversion equation in this paper. Actually, unlike Hotz Miller (1993) or Artuc Chaudhuri McLaren (2010), we never take logarithm of probabilities in the estimation algorithm. Different from the CCP representation, we estimate expected values directly from count data which makes our method more convenient when estimation of probabilities or evaluation of the likelihood function is difficult. This is usually the case when the number of choices and structural parameters is large.

In the next section, we describe the estimation procedure of the generic model we present here. (5), (6) and (7) play key roles in the estimation procedure.

2 Estimation

Our method has two stages: First, the Poisson regression stage, where we estimate expected values associated with each choice for every time period. Second, the Bellman equation

[6]See Appendix B for derivation of the equations.

stage, where we plug estimated expected values into a Bellman equation to construct a linear regression and retrieve structural parameters of the model.

Step 1: PPML Regression

In this step, our goal is to estimate expected values $\tilde{V}_t^{i,s}$ and bilateral resistance parameters $C_t^{ij,s}$. We construct a simple expression for flows between options, similar to the gravity equation, which is essentially a Poisson pseudo maximum likelihood regression available in many different types of statistical software.

The Stage 1 regression equation is

$$y_t^{ij,s} = \exp\left[\Lambda_t^{j,s} + \Gamma_t^{i,s} + \Psi_t^{ij,s}\right] + \xi_t^{1,ij,s}, \tag{8}$$

where $y_t^{ij,s}$ is total number of agents with state (i,s) who choose j (hence $y_t^{ij,s} = L_t^{i,s} m_t^{ij,s}$), $\Lambda_t^{j,s}$ is the destination fixed effect, $\Gamma_t^{i,s}$ is the origin fixed effect, and $\Psi_t^{ij,s}$ is the bilateral resistance term. The equation above can be interpreted as a Poisson pseudo-maximum likelihood regression[7].

Derivation of the Stage 1 regression equation:

If we multiply (6) with L_t^i, we get

$$y_t^{i,s} = \exp\left\{\frac{\beta}{\nu}\tilde{V}_{t+1}^{j,s} - \frac{\beta}{\nu}\tilde{V}_{t+1}^{i,s} + \log\left(L_t^{i,s}\right) - \frac{1}{\nu}\Omega_t^{i,s} - \frac{1}{\nu}C_t^{ij,s}\right\},$$

then we can arrange the terms as i-specific terms, j-specific terms and bilateral terms. (Note that, we need to drop either destination or fixed effect for one choice. Otherwise the regression matrix becomes singular. Assume that we drop the destination fixed effect for the choice $i = 1$).

[7]See Gourieroux, Monfort and Trognon (1984) and Cameron and Trivedi (1998) for properties of the PPML regression.

Then the j-specific term or the destination fixed effect $\Lambda_t^{j,s}$ is equal to

$$\Lambda_t^{j,s} = \frac{\beta}{\nu}\tilde{V}_{t+1}^{j,s} - \frac{\beta}{\nu}\tilde{V}_{t+1}^{1,s},$$

the i-specific term or the origin fixed effect $\Gamma_t^{i,s}$ is equal to

$$\Gamma_t^{i,s} = -\frac{\beta}{\nu}\tilde{V}_{t+1}^{i,s} - \frac{1}{\nu}\Omega_t^{i,s} + \log(L_t^{i,s}) + \frac{\beta}{\nu}\tilde{V}_{t+1}^{1,s},$$

and the bilateral resistance term $\Psi_t^{ij,s}$ is equal to

$$\Psi_t^{ij,s} = -\frac{1}{\nu}C_t^{ij,s}.$$

Note that the option value term $\Omega_t^{i,s}$ can be expressed as

$$\frac{1}{\nu}\Omega_t^{i,s} = -\Lambda_t^{i,s} - \Gamma_t^{i,s} + \log(L_t^{i,s}), \tag{9}$$

Stage 2: Bellman Equation

In Stage 1, we have estimated the expected values, $\Lambda_t^{j,s}$, and moving cost parameters, $\Psi_t^{ij,s}$. The next step is to estimate other parameters, including $1/\nu$. In Stage 2, the goal is to construct the Bellman equation using the estimated parameters from Stage 1 and estimate the remaining parameters.

The Stage 2 regression equation is

$$\phi_t^{i,s} \quad = \quad \zeta_t^s + \widetilde{\eta}^{i,s} + \frac{\beta}{\nu}\widetilde{w}_{t+1}^{i,s} + \xi_t^{i,s}, \tag{10}$$

where ζ_t^s is the time dummy specific to type s, $\widetilde{\eta}^{i,s}$ is the sector dummy specific to s, $\widetilde{w}_{t+1}^{i,s}$ is the expected wage constructed using (12), $\xi_t^{i,s}$ is the regression residual and finally $\phi_t^{i,s}$ is the dependent variable constructed from Step 1 estimates using equation

$$\phi_t^{i,s} \quad = \quad \Lambda_t^{i,s} + \sum_{s' \in S} \pi\left(s, s'\right)\left(\beta \Gamma_{t+1}^{i,s'} - \log(L_{t+1}^{i,s'})\right). \tag{11}$$

The expected wages in (10) are equal to

$$\widetilde{w}_{t+1}^{i,s} \quad = \quad \sum_{s' \in S} \pi\left(s, s'\right) w_{t+1}^{i,s'}. \tag{12}$$

It is possible to use Generalized Method of Moments or Instrumental Variables method for the regression [8].

Derivation of the Stage 2 regression equation:

After multiplying (5) with β/ν, aggregating it over possible states and moving all terms to the left hand side, we get

$$E_t\left[\frac{\beta}{\nu}\widetilde{V}_{t+1}^{i,s} - \frac{\beta}{\nu}\sum_{s' \in S} \pi\left(s, s'\right)\left(w_{t+1}^{i,s'} + \eta^{i,s'} + \beta \widetilde{V}_{t+2}^{i,s'} + \Omega_{t+1}^{i,s'}\right)\right] = 0,$$

Then

$$E_t\left[\Lambda_t^{i,s} - \frac{\beta}{\nu}\widetilde{V}_{t+1}^{1,s} - \frac{\beta}{\nu}\widetilde{w}_{t+1}^{i,s} - \widetilde{\eta}^{i,s} - \sum_{s' \in S} \pi\left(s, s'\right)\left(-\beta \Gamma_{t+1}^{i,s'} + \log(L_{t+1}^{i,s'}) - \frac{\beta^2}{\nu}\widetilde{V}_{t+2}^{1,s'}\right)\right] = 0, \tag{13}$$

where $\widetilde{w}_{t+1}^{i,s}$ is defined in (12) and

[8] For a detailed analysis of identification problems in discrete dynamic models, see Magnac and Thesmar (2002).

$$\widetilde{\eta}^{i,s} = \frac{\beta}{\nu}\sum_{s'\in S}\pi\left(s,s'\right)\eta^{i,s'}.$$

We define

$$\zeta_t^s = \frac{\beta}{\nu}\widetilde{V}_{t+1}^{1,s} + \sum_{s'\in S}\pi\left(s,s'\right)\left(\frac{\beta^2}{\nu}\widetilde{V}_{t+2}^{1,s'}\right),$$

then we can re-arrange (13) and write it as

$$E_t\left[\phi_t^{i,s} - \zeta_t^s - \widetilde{\eta}^{i,s} - \frac{\beta}{\nu}\widetilde{w}_{t+1}^{i,s}\right] = 0.$$

Alternative specifications:

We focus on models that can be estimated using repeated cross-section data with retrospective questions, such as household labor force surveys which are available for many countries. An example from the US is the March supplement of Current Population Survey[9]. However if longitudinal data are available, it is possible to consider unobserved heterogeneity in the model. Arcidiacono and Miller (2011) show how an EM loop can be incorporated in CCP to estimate unobserved heterogeneity. Their intuition can also be applied to PPML regression. Appendix D illustrates how it is possible to use an EM loop within PPML regression when panel data are available.

Another alternative modeling approach is to use wage shocks rather than moving cost shocks in agents' utility function. In Appendix A, we provide an equation that can be used instead of (10) in case of such wage shocks.

In the next section we present an example to illustrate a practical application of the method.

[9]Other countries with such data, that we are aware of so far, are Indonesia, Mexico and Turkey.

3 Example Application: Sectoral Mobility in the US

In this section we present an application of the estimation method. First, we estimate a disaggregated variant of Artuc, Chaudhuri and McLaren (2010) using the exact same data, which is the Current Population Survey from years 1975 to 2001 (henceforth CPS). Recently, Kaplan, Lederman and Robertson (2013), Artuc and McLaren (2012) and Artuc, Bet, Brambilla and Porto (2013) used the estimation method we introduce herein.

Model

To elaborate on the generic model we presented in the previous section, consider that sectors are industries in which workers choose to work in each time period. For each choice, workers receive a payoff w_t^i and an idiosyncratic utility η^i common to all workers in sector i. Assume that $\eta^1 = 0$ for normalization. For simplicity, we consider one type of worker, hence drop the state superscript s. We allow the deterministic moving cost to change over time such that $C_t^{ij} = c_t$ if $i \neq j$ and $C_t^{ij} = 0$ if $i = j$.

We use two regressions to estimate structural parameters of the model. First, the Poisson regression equation is

$$y_t^{ij} = \exp\left(\Gamma_t^i + \Lambda_t^j + \Psi_t 1_{i\neq j}\right) + e_t^{ij}, \tag{14}$$

where the regression coefficient $\Psi_t = -c_t/\nu$ and the indicator function $1_{i\neq j}$ is equal to one when $i \neq j$ and zero otherwise, and e_t^{ij} is the residual. In many cases, the discount rate can not be identified, therefore we assume that it is equal to $\beta = 0.97$, and known by the econometrician.

Second, the regression equation based on the Bellman equation is

$$\phi_t^i = \zeta_t + \widetilde{\eta}^i + \frac{\beta}{\nu} w_{t+1}^i + \xi_t^i, \tag{15}$$

where ξ_t^i is the residual, $\phi_t^i = \Lambda_t^i + \beta\Gamma_{t+1}^i - \log(L_{t+1}^i)$ is the dependent variable, $\widetilde{\eta}^i$ is a sector dummy, and ζ_t is a time dummy. We set $\widetilde{\eta}^1 = 0$ for the first sector.

In the "Alternative specification" we use different set of sector dummies, we allow the sector specific fixed utility to have linear time trends. This specification potentially captures changes in employment opportunities in different sectors over time. Modifying the assumption on fixed utility only affects the second stage regression.

The second stage regression for the "Alternative Specification" is

$$\phi_t^i = \zeta_t + \widetilde{\eta}_1^i + \widetilde{\eta}_2^i t + \frac{\beta}{\nu} w_{t+1}^i + \xi_t^i, \tag{16}$$

where $\widetilde{\eta}_1^i$ is the intercenp and $\widetilde{\eta}_2^i$ is the trend. We set $\widetilde{\eta}_1^1 = 0$ and $\widetilde{\eta}_2^1 = 0$ for the first sector.

Data

Artuc, Chaudhuri and McLaren (2010) use CPS data for males between 25 to 64 years old, from the year 1976 to 2001; we use the same data and sample section procedure. CPS is a repeated cross-section, and its March supplement provides retrospective industry questions regarding workers' industry in the previous year, along with their current industry. These retrospective questions allow us to construct number of workers moving from industry i to j, denoted as y_t^{ij}. In addition to flow data, we use average wage data for each industry. Artuc, Chaudhuri and McLaren (2010) aggregate industries to 6 major sectors[10]. Different from them, we aggregate industries to 16 sectors. The sectors are: 1. Agriculture, 2. Mining, 3. Construction, 4. Non-durable manufacturing, 5. Durable manufacturing, 6. Transportation, 7. Communications, 8. Utilities, 9. Wholesale trade, 10. Retail trade, 11. Finance, 12. Business, 13. Personal services, 14. Entertainment, 15. Professional, and 16. Public.

In addition to increasing the number of choices, we consider sector specific iid utility shocks, η^i, and let the deterministic part of moving cost, c_t, change over time. These two changes improve their theoretical model significantly because some sectors may be more preferable by workers for non-pecuniary reasons and the moving costs may change over the twenty six year sample. These possibilities are now addressed in the model. We use PPML regression in the first step and IV regression in the second step. Everything else is exactly

[10] They were not able to disaggregate sectors further because their method did not allowed zero cells in the transition matrix. The were able to estimate at most seven structural parameters.

the same as their basic model and benchmark regression, including the choice of instruments.

Results

We estimate the distributional parameter $1/\nu$, 15 parameters for η^i, and 26 parameters for c_t, thus 42 structural parameters total. In the first stage, we estimate c_t/ν, and destination and origin fixed effects using equation (14). Then, we construct the second stage regression equation (16) using the destination and origin fixed effects from the first stage regression. In the second stage, we estimate the remaining structural parameters, η^i_t/ν and $1/\nu$. We use a one year lag for the second stage IV regression.

Table 1 shows the estimation results for the basic specification. We present robust standard errors in the first stage regression. In the first step, all coefficients are significant at 1 percent level. We find that C_t/ν changes between 4.49 and 4.88, with an average of 4.67.

In the second step, $1/\nu$ is estimated as 0.96 and is significant at 1 percent level, and 9 out of 15 unobserved utility coefficients are significant at 1 percent level and 5 coefficients are significant at 5 percent level.

Table 2 shows the estimation results for the alternative specification. The first stage regression for the alternative specification is identical to the basic specification, thus C_t/ν estimates are exactly the same. We find that $1/\nu$ is estimated as 3.67 which is much larger than the basic specification estimate. (Note that larger $1/nu$ means smaller ν and C).

In the following section, we simulate data for steady state and transition under policy shocks and re-estimate the model with simulated data to illustrate performance of our estimation method relative to other methods in the literature.

4 Monte Carlo Simulations

Running counter-factual policy simulations is usually the main motivation for structural estimation. Reduced form equations are subject to Lucas critique and can not be used in policy simulations[11]. Although we use estimators which are traditionally reduced form, each coefficient in the regression equations corresponds to a structural parameter. Using the structural parameters, it is possible to simulate the model presented in the previous sections under different policy scenarios. However, in this paper, we are not interested in particular effects of policies per se: Our goal is to show the performance of this new estimation method using simulated data. We expose the system to policy shocks and illustrate robustness of the estimation method under non-stationary conditions. In a sense, we create aggregate shocks artificially.

For an illustration, we consider an open economy model with trade shocks, exogenous prices and endogenous wages. To simulate trade shocks, we need to define equilibrium real wages as functions of labor supply and prices. Assume that sectors are perfectly competitive with simple Cobb-Douglas production functions. We assume that workers are paid their real marginal products. Then, the following real wage equation closes the model

$$w_t^i = (p_t^i/P_t)a_i\tilde{A}^i(L_t^i)^{a_i-1}, \tag{17}$$

where p_t^i is the exogenous price of sector i output, P_t is the consumer price index $P_t = \Pi_i(p_t^i)^{b_i}$ with basket shares b_i, and $\tilde{A}^i$ is a constant that is calibrated from the data.

We calculate Cobb-Douglas labor shares and consumer basket shares from Bureau of Economic Analysis data. Then, we calibrate $\tilde{A}^i$ to match average wages in given sectors. The calibration exercise is similar to Artuc Chaudhuri McLaren (2010). The production function and consumer price index parameters are reported in Table 2 along with wages and labor allocations. We normalize all prices to one at steady state, $p_t^i = 1$ for $i = 1,..,16$, and fix the deterministic part of moving cost to be constant[12] over time $c_t = 4.5$. We assume

[11] As an example, consider a policy experiment of reducing the moving costs, C_t, by 50 per cent. Assume that we would like to know the effect of this change on workers' mobility decisions. We cannot simply change the resistance coefficient Ψ_t^{ij} and keep other coefficients as they were. Because after a change in the moving cost, the values would also change, thus the Γ_t^i and Λ_t^i parameters would change as well. So, it is impossible to use reduced form parameters Γ_t^i and Λ_t^i for simulations. Because of Lucas critique, one has to know the underlying structural parameters.

[12] We assume that moving costs are constant over time for cosmetic reasons, so that the results are easy

$\nu = 1$ and assign arbitrary values to η^i.

For the simulations, we use a multiple shooting algorithm similar to Lipton et al (1982), but one can use other shooting methods instead.

We consider four simulation exercises:

In Simulation I, we simulate the model around steady state[13]. Then, we estimate the model using 26 years of simulated data.

In Simulation II, we drop the manufacturing prices 20 percent as a surprise one time shock, which implies a tariff reduction in the protected manufacturing industries (sectors 4 and 5). After this one time shock, we let the system reach new steady state over time. Then, we estimate the model using simulated data during this transitory period.

In Simulation III, we increase the number of years from 26 to 100 to show the asymptotic properties of the estimation method.

In Simulation IV, we decrease the number of choices from 16 to 8 to show the impact of having a smaller number of observations because of smaller number of choices.

Then, we repeat all four simulation exercises 300 times. All simulations are conducted with $L = 20,000$ agents, which is approximately equal to the sample size of March-CPS that is used for the estimation in the previous section.

Table 4 presents the Monte Carlo simulation results. The column labeled as "Sim I" shows that the estimates are reasonably close to the true values and expected to be unbiased.

The column "Sim II" shows that using data contaminated with a non-stationary trade policy shock does not affect the performance of the method. Note that we did not specify the nature of the aggregate shock in the estimation procedure. The method introduced herein does not require strong distributional assumptions about the aggregate shocks. CCP method and other maximum likelihood based methods require the aggregate shocks to be

to read.

[13]Note that wages show some fluctuations over time, for that reason we added an iid normal shock to equilibrium wages with standard deviation equal to 0.05, approximately equal to the standard error of average wages in the data. We also added an unexpected surprise shock to the wages with a standard deviation equal to 0.05.

fully specified and to be stationary.

The column "Sim III" presents the results for the longer time series with 100 years. It hints that the method has plausible asymptotic properties, i.e. standard errors decrease as we increase the length of time series, and the estimates converge to the true parameter values.

Finally, column "Sim IV" shows that as the number of choices decrease, the standard errors increase in both stages.

In the following tables, 5 and 6, we compare results of PPML and CCP based estimation strategies. We use "Simulation I" data with $20,000$ agents, then we repeat the exercise with $2,000$ and $4,000$ agents to demonstrate small sample properties of the estimators.

Table 5 shows estimated values and standard errors in parentheses, Λ^i, using PPML and CCP methods with $2,000$ agents, with $4,000$ agents, with $20,000$ agents and with infinitely many agents[14]. We use equal weights in the non-parametric stage of CCP estimation. The last two columns show that both PPML and CCP methods converge to the true values, therefore they are asymptotically equivalent. With finite number of agents, PPML estimates are closer to the true values, especially when the sample size is small. However, we do not argue that PPML is more efficient than CCP, because the non-parametric stage of CCP estimation could be conducted with different weights and we cannot try all possible weighting vectors. When the number of choices are large, PPML is more convenient than non-parametric CCP estimation since it does not rely on taking logarithms of probabilities that can be very close to zero.

Table 6 presents the estimation results (and standard errors in parentheses) for C/ν and $1/\nu$ using different estimation methods. When we use a maximum likelihood based method, we assume that the econometrician knows the distribution of aggregate shocks to wages in order to use maximum likelihood estimation. Also we assume that the econometrican knows the true values of η's, because otherwise repeating ML estimation procedure 300 times takes

[14] The econometrician can observe true switching probabilities m_t^{ij} when there are infinitely many agents. But there may still be uncertainty due to the aggregate shocks to w_t^i.

unreasonably long time.

PPML1 is the method described in the previous section, which is the two stage procedure with PPML estimation in the first stage and linear regression in the second stage. PPML2 method uses PPML to impute expected values within a maximum likelihood estimation algorithm. It is similar to CCP, but rather than imputing expected values non-parametrically we use PPML. ACM method is the estimator used in Artuc, Chaudhuri and McLaren (2010)[15]. CCP method is the conditional choice probability method that uses maximum likelihood and non-parametric estimation of expected values with Hotz-Miller inversion equation. In the CCP and PPML2 methods, we assume that the exonometrician knows the distribution of aggregate shocks, since it is needed for the maximum likelihood estimation. Table 6 shows that all four methods perform well with large sample. However, CCP method did not converge when the sample size was small ($L = 2,000$). PPML based methods seem to perform better when sample size is small, also PPML1 has an important advantage over ML-based methods since it does not require distributional assumptions about the aggregate shocks.

In the next Monte-Carlo exercise, we shut down the aggregate shocks to wages. Without aggregate shocks, it is straightforward use iterative methods pioneered by Rust (1987). Iterative estimation methods are out of the scope of this paper but the "Nested Psudo-Maximum Likelihood" method introduced by Aguirregabiria and Mira (2002) is relevant and important. They showed that it is possible to use an iterative step to improve the performance of CCP with small samples. Without aggregate shocks, we are able to compare performance of the NPM with PPML methods.

Table 6 presents results of Monte-Carlo simulations with PPML1, PPML2, ACM CCP and NPM (standard errors are in parentheses). With all five methods, estimates converge to the true parameter values as the sample size increases. NPM indeed improves the CCP results for small samples, however it is very difficult to implement it when there are aggregate

[15] Unlike Artuc Chaudhuri and McLaren (2010), we have zero cells in the transition matrix where $m_t^{ij} = 0$ for some i, j, t because we have 16 choices rather than 6. Different from them, we drop the observation when $m_t^{ij} = 0$, which makes the ACM estimator biased for this particular exercise.

shocks. CCP cannot be used as starting point for the NPM algorithm when the sample size is very small (when we simulated 2000 agents the CCP method did not converge to finite numbers). Naturally, it is possible to use PPML estimates as a starting point for the NPM algorithm. The row labeled as "PPML-NPM" shows estimates of a variation of NPM method that uses PPML as a starting point rather than CCP. The extra NPM-loop after the PPML regression reduces the standard errors, but it is difficult to implement when there are aggregate shocks.

5 Conclusion

We present a novel and computationally efficient method for estimating dynamic discrete choice models with heterogeneity and time-varying resistance (i.e moving cost) parameters. The method performs well with large number of choices, sparse decision transition matrices (caused by small sample size) and aggregate shocks. All expectations of agents are fully accounted for in the first step regression, which allows us to be agnostic about agents' expectations and distribution of aggregate shocks. Therefore the method can be used for estimation out of steady state. Potential applications are migration, sectoral and occupational labor mobility models with large number of discrete choices, macroeconomic shocks and limited heterogeneity.

Appendix A: An Alternative Model with Wage Shocks

The moving cost shock ε_t^i, in essence, is a utility shock. However, it is common in the labor economics literature to consider wage shocks, rather than utility shocks, as the main driving force behind labor mobility. Consider an alternative specification where ε_{t-1}^i is a wage shock that is revealed at the end of time $t-1$ but affects observed wage at time t, rather than a utility shock. Assume that the econometrician observes $\bar{w}_t^i = w_t^i + \varepsilon_{t-1}^i$, but not w_t^i. Then (10) can not be used as the basis for the regression, since observed wages are self-selected and $\bar{w}_t^i$ is a biased measure of true underlying sectoral wage w_t^i.

With wage shocks, there are not any changes in the PPML regression step, but the second step has to be modified. The expected wage conditional on being in sector i is equal to

$$E_t\left(w_t^i + \varepsilon_{t-1}^i | i\right) = w_t^i - \nu \sum_{j=1}^{n} \widetilde{m}_t^{ji} \log\left(m_t^{ji}\right),$$

where $\widetilde{m}_t^{ji}$ is the ratio of agents who switch from j to i conditional on being in sector i in period t,

$$\widetilde{m}_t^{ji,s} = \frac{L_t^{j,s} m_t^{ji,s}}{\sum_{k=1}^{n} L_t^{k,s} m_t^{ki,s}}.$$

Assume that

$$\mu_t^{i,s} = -\sum_{j=1}^{n} \widetilde{m}_t^{ji,s} \log\left(m_t^{ji,s}\right),$$

then the wages in the second stage regression equation should be corrected using $\mu_t^{i,s}$. Derivation of the equations are provided in Appendix B.3.

Appendix B: Derivation of Key Equations

As noted in the main text, the cdf for the extreme value type I distribution with location parameter $-\nu\gamma$ and scale parameter ν is :

$$F(\varepsilon) = \exp(-\exp(-\varepsilon/\nu - \gamma)),$$

where $E(\varepsilon) = 0$, $Var(\varepsilon) = \pi^2\nu^2/6$ and γ is the Euler's constant ($\gamma \cong 0.577$). Then, pdf is:

$$f(\varepsilon) = (1/\nu)\exp(-\varepsilon/\nu - \gamma - \exp(-\varepsilon/\nu - \gamma)).$$

B.1 Gross Flow Function

We are dropping the state superscript s and time subscript t for notational convenience. Define

$$\overline{\varepsilon}^{ij} = [\beta\widetilde{V}^j - \beta\widetilde{V}^i] - C^{ij},$$

The gross flow function, m^{ij}, is equal to the probability that a given i sector worker will switch to j sector, that is the probability of a sector i worker to have higher utility in sector j in the next period. This probability is

$$m^{ij} = \Pr\left[\overline{\varepsilon}^{ij} + \varepsilon^j \geq \overline{\varepsilon}^{ik} + \varepsilon^k \text{ for } k = 1, \ldots, n\right],$$

this can be written as

$$m^{ij} = \int_{-\infty}^{\infty} f(\varepsilon^j) \prod_{k \neq j} F(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}) d\varepsilon^j.$$

Thanks to the extreme value distribution and McFadden (1973), the gross flow $m_t^{ij,s}$ can be written as

$$m_t^{ij} = \frac{\exp\left(\left(\beta\widetilde{V}_{t+1}^j - \beta\widetilde{V}_{t+1}^i - C_t^{ij}\right)\frac{1}{\nu}\right)}{\sum_{k=1}^{N}\exp\left(\left(\beta\widetilde{V}_{t+1}^j - \beta\widetilde{V}_{t+1}^i - C_t^{ij}\right)\frac{1}{\nu}\right)},$$

which is equal to

$$m_t^{ij} = \frac{\exp\left(\overline{\varepsilon}_t^{ij}/\nu\right)}{\sum_{k=1}^{N}\exp\left(\overline{\varepsilon}_t^{ik}/\nu\right)}.$$

B.2 Option Values

We follow the steps in Artuc, Chaudhuri and McLaren (2010). Define, for convenience: $x = \varepsilon^j/\nu + \gamma$ and $z = \log(\frac{\sum_{k=1}^{n}\exp(z^k/\nu)}{\exp(z^k/\nu)})$.

Now, define:

$$\begin{aligned}
\Phi^{ij} &\equiv \textstyle\int_{-\infty}^{\infty} \varepsilon^j f(\varepsilon^j) \prod_{j\neq k} F(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}) d\varepsilon^j \\
&= \tfrac{1}{\nu} \textstyle\int \varepsilon^j \exp(-\varepsilon^j/\nu - \gamma - \exp(-\varepsilon^j/\nu - \gamma)) \prod_{k\neq j} \exp(-\exp(-[\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}]/\nu - \gamma)) d\varepsilon^j
\end{aligned}$$

Then,

$$\begin{aligned}
\Phi^{ij} &= \tfrac{1}{\nu} \textstyle\int \varepsilon^j \exp(-\varepsilon^j/\nu - \gamma - \exp(-\varepsilon^j/\nu - \gamma)) \exp(-\sum_{k\neq j} \exp(-[\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}]/\nu - \gamma)) d\varepsilon^j \\
&= \tfrac{1}{\nu} \textstyle\int \varepsilon^j \exp(-\varepsilon^j/\nu - \gamma) \exp(-\sum_{k=1}^{n} \exp(-[\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}]/\nu - \gamma)) d\varepsilon^j \\
&= \tfrac{1}{\nu} \textstyle\int \varepsilon^j \exp\left[(-\varepsilon^j/\nu - \gamma) - \sum_{k=1}^{n} \exp(-[\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}]/\nu - \gamma)\right] d\varepsilon^j \\
&= \tfrac{1}{\nu} \textstyle\int \varepsilon^j \exp\left[(-\varepsilon^j/\nu - \gamma) - \exp((-\varepsilon^j/\nu - \gamma)) \sum_{k=1}^{n} \exp(-[z^j - z^k]/\nu)\right] d\varepsilon^j \\
&= \tfrac{1}{\nu} \textstyle\int \varepsilon^j \exp\left[(-\varepsilon^j/\nu - \gamma) - \exp((-\varepsilon^j/\nu - \gamma)) \left(\sum_{k=1}^{n} \exp(z^k/\nu)\right) / \exp(z^j/\nu)\right] d\varepsilon^j
\end{aligned}$$

Note that, $x = \varepsilon^j/\nu + \gamma$ and $z = \log(\frac{\sum_{k=1}^{n}\exp(z^k/\nu)}{\exp(z^k/\nu)})$. Then,

$$
\begin{aligned}
\Phi^{ij} &= \int \varepsilon^j \exp(-x - \exp(-(x-z)))dx \\
&= \int \nu(x-\gamma) \exp(-x - \exp(-(x-z)))dx \\
&= (-\nu\gamma) \exp(-z) + \nu \int x \exp(-x - \exp(-(x-z)))dx \\
&= (-\nu\gamma) \exp(-z) + \nu \exp(-z) \int x \exp(-x + z - \exp(-(x-z)))dx
\end{aligned}
$$

We know that $\exp(-z) = m^{ij}$ from McFadden (1973). Substituting this in:

$$
\begin{aligned}
\Phi^{ij} &= (-\nu\gamma)m^{ij} + \nu m^{ij} \int x \exp(-x + z - \exp(-(x-z)))dx \\
&= (-\nu\gamma)m^{ij} + \nu m^{ij} \int x \exp(-x + z - \exp(-(x-z)))dx \\
&\quad + \nu m^{ij} \int z \exp(-x + z - \exp(-(x-z)))dx \\
&\quad - \nu m^{ij} \int z \exp(-x + z - \exp(-(x-z)))dx
\end{aligned}
$$

Then we set $y = x - z$, thus

$$
\begin{aligned}
\Phi^{ij} &= (-\nu\gamma)m^{ij} + \nu m^{ij} \int (x-z) \exp(-x + z - \exp(-(x-z)))dx \\
&\quad + \nu m^{ij} \int z \exp(-x + z - \exp(-(x-z)))dx
\end{aligned}
$$

$$\begin{aligned}\Phi^{ij} &= (-\nu\gamma)m^{ij} + \nu m^{ij}\int y\exp(-y-\exp(-y))dy + \nu z m^{ij}\int \exp(-y-\exp(-y))dy \\ &= (-\nu\gamma)m^{ij} + \nu m^{ij}\int y\exp(-y-\exp(-y))dy + \nu z m^{ij}.\end{aligned}$$

Noting that $\int y\exp(-y-\exp(-y))dy = \gamma$ (Euler's constant), we can simplify:

$$\begin{aligned}\Phi^{ij} &= (-\nu\gamma)m^{ij} + \nu z m^{ij} + \nu\gamma m^{ij} \\ &= -\nu\log(m^{ij})m^{ij}\end{aligned}$$

Then we can add this across possible destinations j, note that the utility of a worker in i is equal to:

$$V_t^i = u_t^i + \sum_{j=1}^{n}\left(\Phi_t^{ij} - m_t^{ij}C_t^{ij} + \beta m_t^{ij}\widetilde{V}_{t+1}^j\right)$$

$$= u_t^i + \sum_{j=1}^{n} m_t^{ij}\left[-\nu\log(m_t^{ij}) - C_t^{ij} + \beta\widetilde{V}_{t+1}^j)\right]$$

$$= u_t^i + \sum_{j=1}^{n} m_t^{ij}\left[-\nu\log(m_t^{ij}) - C_t^{ij} + \beta(\widetilde{V}_{t+1}^j - \widetilde{V}_{t+1}^i)\right] + \beta\widetilde{V}_{t+1}^i$$

$$= u_t^i + \sum_{j=1}^{n} m_t^{ij}\left[\overline{\varepsilon}_t^{ij} - \nu\log(m_t^{ij})\right] + \beta\widetilde{V}_{t+1}^i.$$

Now, recall from above that $\log(m^{ij}) = \bar{\varepsilon}_t^{ij}/\nu - \log\left(\sum_{k=1}^n \exp(\bar{\varepsilon}^{ik}/\nu)\right)$. This yields:

$$\begin{aligned} V_t^i &= u_t^i + \sum_{j=1}^n m_t^{ij}\left[\nu \log\left(\sum_{k=1}^n \exp(\bar{\varepsilon}^{ik}/\nu)\right)\right] + \beta \widetilde{V}_{t+1}^i \\ &= u_t^i + \nu \log\left(\sum_{k=1}^n \exp(\bar{\varepsilon}^{ik}/\nu)\right) + \beta \widetilde{V}_{t+1}^i. \end{aligned}$$

This implies that the option value Ω^i can be written as

$$\Omega^i = -\nu \log\left(m^{ii}\right).$$

B.3 Wage Shocks

Assume that d_t^i denotes agent's choice at time t. Expected ε^j conditional on a sector i agent choosing sector i is equal to

$$\begin{aligned} E\left(\varepsilon_t^i | d_t = i, d_{t+1} = j\right) &= \frac{\int_{-\infty}^{\infty} \varepsilon^j f(\varepsilon^j) \prod_{j \neq k} F(\varepsilon^j + \overline{\varepsilon_t}^{ij} - \overline{\varepsilon_t}^{ik}) d\varepsilon^j}{\int_{-\infty}^{\infty} f(\varepsilon^j) \prod_{j \neq k} F(\varepsilon^j + \overline{\varepsilon_t}^{ij} - \overline{\varepsilon_t}^{ik}) d\varepsilon^j} \\ &= \frac{\Phi_t^{ij}}{m_t^{ij}} \\ &= -\nu \log\left(m_t^{ij}\right) \end{aligned}$$

Adding this across possible origins, we find

$$E_t\left(\varepsilon_{t-1}^j | d_t = j\right) = -\nu \sum_{i=1}^n \widetilde{m}_t^{ij} \log\left(m_t^{ij}\right),$$

where $\widetilde{m}_t^{ji}$ is the probability of a sector i agent to originate from sector j

$$\widetilde{m}_t^{ij} = \frac{L_t^{i,s} m_t^{ij}}{\sum_{k=1}^n L_t^k m_t^{kj}}.$$

Appendix C: CCP Representation of the Model

Following the steps in Appendix B.2 and using (3) the Bellman equation can be written as

$$
\begin{aligned}
V_t^{i,s} &= u_t^{i,s} + \max_j \left\{ \beta \widetilde{V}_{t+1}^{j,s} - C_t^{ij,s} - \varepsilon_t^j \right\}, \\
&= u_t^{i,s} + \sum_j m_t^{ij} \left(\beta \widetilde{V}_{t+1}^{j,s} - C_t^{ij,s} + \frac{\Phi_t^{ij,s}}{m_t^{ij,s}} \right), \\
&= u_t^{i,s} + \sum_j m_t^{ij} \left[-\nu \log(m_t^{ij}) - C_t^{ij} + \beta(\widetilde{V}_{t+1}^{j,s} - \widetilde{V}_{t+1}^{k,s})) \right] + \beta \widetilde{V}_{t+1}^{k,s},
\end{aligned}
$$

$$
\begin{aligned}
&= u_t^{i,s} + \sum_j m_t^{ij} \left[C_t^{ij} - \beta(\widetilde{V}_{t+1}^{j,s} - \widetilde{V}_{t+1}^{i,s}) + \nu \log \left(\sum_{n=1}^N \exp\left(\bar{\varepsilon}_t^{in}/\nu\right) \right) - C_t^{ij} + \beta(\widetilde{V}_{t+1}^{j,s} - \widetilde{V}_{t+1}^{k,s}) \right] + \beta \widetilde{V}_{t+1}^{k,s} \\
&= u_t^{i,s} + \sum_j m_t^{ij} \left[\beta \widetilde{V}_{t+1}^{i} + \nu \log \left(\sum_{n=1}^N \exp\left(\bar{\varepsilon}_t^{in}/\nu\right) \right) - \beta \widetilde{V}_{t+1}^{k,s} \right] + \beta \widetilde{V}_{t+1}^{k,s} \\
&= u_t^{i,s} + \sum_j m_t^{ij} \left[\beta \widetilde{V}_{t+1}^{i} + \nu \log \left(\sum_{n=1}^N \exp\left(\bar{\varepsilon}_t^{in}/\nu\right) \right) - \beta \widetilde{V}_{t+1}^{k,s} + C_t^{ik,s} \right] + \beta \widetilde{V}_{t+1}^{k,s} - C_t^{ik,s} \\
&= u_t^{i,s} + \beta \widetilde{V}_{t+1}^{k,s} - C_t^{ik,s} - \nu \log \left(m_t^{ik,s} \right)
\end{aligned}
$$

Then

$$
\begin{aligned}
V_t^{j,s} - V_t^{i,s} &= u_t^{j,s} + \beta \widetilde{V}_{t+1}^{k,s} - C_t^{jk,s} - \nu \log \left(m_t^{jk,s} \right) - u_t^{i,s} - \beta \widetilde{V}_{t+1}^{k,s} + C_t^{ik,s} + \nu \log \left(m_t^{ik,s} \right), \\
&= (u_t^{j,s} - u_t^{i,s}) - C_t^{jk,s} - \nu \log \left(m_t^{jk,s} \right) + C_t^{ik,s} + \nu \log \left(m_t^{ik,s} \right),
\end{aligned}
$$

since k can be any sector, we can add the expression over all possible sectors to increase precision

$$V_t^{j,s} - V_t^{i,s} = (u_t^{j,s} - u_t^{i,s}) + \sum_{k=1} x^{ij,k}\left(-C_t^{jk,s} - \nu \log\left(m_t^{jk,s}\right) + C_t^{ik,s} + \nu \log\left(m_t^{ik,s}\right)\right), \quad (18)$$

where $x_t^{ij,k}$ is an arbitrary weighting vector such that $\sum_k x_t^{ij,k} = 1$.

Then (18) is the CCP representation of the model. Note that $u_t^{i,s} = w_t^{i,s} + \eta^{i,s}$. We need guessed values of $\eta^{i,s}$ for the CCP representation of the model. Therefore, we need to estimate all parameters at once if we use CCP and maximum likelihood. Which makes CCP computationally demanding when the number of choices and structural parameters are large. Also may be difficult to use CCP in certain cases when many of the observed conditional choice probabilities are close to zero.

Appendix D: EM loop within PPML regression

It is possible to incorporate Expectation-Maximization algorithm to our estimation procedure in the first step. For notational convenience we consider the case where agents' current and last two sectors are observed in the data, it is straightforward to generalize this procedure for panels with longer time dimensions.

Assume that we observe each agent's decision at time t and $t+1$, let us denote agent's location at time t with i, time $t+1$ with j, and time $t+2$ with k. The two period flow at time t is denoted with $m_t^{ijk,s}$, the number of workers who chose i, j, and k consecutively is equal to $y_t^{ijk,s} = L_t^i m_t^{ijk,s}$. Each agent has an unobserved discrete type $\sigma \in \tilde{S}$, the observed states (or types) are still denoted with $s \in S$. We are interested in finding the ratio of type σ workers in the observed flow $m_t^{ijk,s}$, let us denote this probability with $\zeta_t^{ijk,s,\sigma}$. Then the number of workers with state (i, s, σ) who choose j then k starting at time t is equal to $\zeta_t^{ijk,s,\sigma} y_t^{ijk,s}$.

Note that

$$\zeta_t^{ijk,s,\sigma} m_t^{ijk,s} = \frac{\exp\left(\left(\beta \widetilde{V}_{t+1}^{j,s,\sigma} - \beta \widetilde{V}_{t+1}^{i,s,\sigma} - C_t^{ij,s,\sigma}\right)\frac{1}{\nu}\right)}{\sum_{n=1}^{N} \exp\left(\left(\beta \widetilde{V}_{t+1}^{n,s,\sigma} - \beta \widetilde{V}_{t+1}^{i,s,\sigma} - C_t^{in,s,\sigma}\right)\frac{1}{\nu}\right)} \cdot \frac{\exp\left(\left(\beta \widetilde{V}_{t+2}^{k,s,\sigma} - \beta \widetilde{V}_{t+2}^{j,s,\sigma} - C_{t+1}^{jk,s,\sigma}\right)\frac{1}{\nu}\right)}{\sum_{n=1}^{N} \exp\left(\left(\beta \widetilde{V}_{t+2}^{n,s} - \beta \widetilde{V}_{t+2}^{j,s} - C_{t+1}^{jn,s,\sigma}\right)\frac{1}{\nu}\right)},$$

which can be represented in log-linear format to construct a Poisson regression as we show in the first step.

Assume that we have an initial guess for $Z_t^{ijk,s,\sigma}$ for time t. Let us denote this initial guess with $Z_t^{ijk,s,\sigma,(1)}$. Using this initial guess, we can run the Poisson regression in the first step

$$\log\left(Z_t^{ijk,s,\sigma,(1)} y_t^{ijk,s}\right) = \Delta_t^{i,s,\sigma,(1)} + \Gamma_t^{j,s,\sigma,(1)} + \Lambda_t^{k,s,\sigma,(1)} + \Psi_t^{ijk,s,\sigma,(1)} X_t^{ijk,s} + e_t^{ijk,s,\sigma}$$

Then the updated probability will be

$$Z_t^{ijk,s,\sigma,(2)} = \frac{\Delta_t^{i,s,\sigma,(1)} + \Gamma_t^{j,s,\sigma,(1)} + \Lambda_t^{k,s,\sigma,(1)} + \Psi_t^{ijk,s,\sigma,(1)} X_t^{ijk,s}}{\sum_{\sigma' \in \tilde{S}} \left(\Delta_t^{i,s,\sigma',(1)} + \Gamma_t^{j,s,\sigma',(1)} + \Lambda_t^{k,s,\sigma',(1)} + \Psi_t^{ij,s,\sigma',(1)} X_t^{ij,s}\right)},$$

Hence guess at step $\tau + 1$ will be

$$Z_t^{ijk,s,\sigma,(\tau+1)} = \frac{\Delta_t^{i,s,\sigma,(\tau)} + \Gamma_t^{j,s,\sigma,(\tau)} + \Lambda_t^{k,s,\sigma,(\tau)} + \Psi_t^{ijk,s,\sigma,(\tau)} X_t^{ijk,s}}{\sum_{\sigma' \in \tilde{S}} \left(\Delta_t^{i,s,\sigma',(\tau)} + \Gamma_t^{j,s,\sigma',(\tau)} + \Lambda_t^{k,s,\sigma',(\tau)} + \Psi_t^{ijk,s,\sigma',(\tau)} X_t^{ijk,s}\right)}.$$

Our simulations confirm that $Z_t^{ijk,s,\sigma,(\tau+1)}$ converges to true $Z_t^{ijk,s,\sigma}$ and we obtain consistent estimates for the destination and origin fixed effects[16].

[16] Results are available upon request.

References

[1] Anderson, James (2010). "The Gravity Model," *The Annual Review of Economics*, 3(1).

[2] Aguirregabiria, Victor and Pedro Mira (2002). "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70(4).

[3] Aguirregabiria, Victor and Pedro Mira (2010). "Dynamic Discrete Choice Models: A Survey," *Journal of Econometrics*, 156.

[4] Arcidiacono, Peter, and Robert Miller (2011). "CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," forthcoming in *Econometrica*.

[5] Artuc, Erhan, Shubham Chaudhuri and John McLaren (2010). "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100(3).

[6] Artuc, Erhan, and John McLaren (2012). "Trade Policy and Wage Inequality: A Structural Analysis with Occupational and Sectoral Mobility," *NBER Working Paper: 18503*.

[7] Artuc, Erhan, German Bet, Irene Brambilla and Guido Porto (2013). "Trade Shocks, Firm Level Investment Inaction and Labor Market Responses," *Mimeo: World Bank*.

[8] Baine, Michel, Frederic Docquier and Caglar Ozden (2009). "Diasporas," *Journal of Development Economics*, 95(1).

[9] Cameron, Colin and Pravin Trivedi (1998). "Regression Analysis of Count Data," Cambridge University Press, Cambridge.

[10] Gourieroux, C., A. Monfort, and A. Trognon (1984). "Pseudo maximum likelihood methods: Applications to Poisson models.," *Econometrica*, 52(3).

[11] Grogger, Jeffrey and Gordon Hanson (2011). "Income Maximisation and the Selection and Sorting of International Migrants.," *Journal of Development Economics*, 95(1).

[12] Hotz, V. Joseph, and Robert Miller (1993). "Conditional Choice Probabilities and the Estimation of Dynamic Models.," *Review of Economic Studies*, 60(3).

[13] Kaplan, D.S., D. Lederman and R. Robertson (2013). "Worker-Level Adjustment Costs in a Developing Country: Evidence from Mexico.," *Mimeo: World Bank*.

[14] Lipton, D., J. Poterba, J. Sachs, L. Summers (1982). "Multiple shooting in rational expectation models," *Econometrica,* 50.

[15] Magnac, Thierry, and David Thesmar (2002). "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70(2).

[16] McFadden, Daniel (1973). "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka (ed.) *Frontiers in Econometrics*, New York, Academic Press.

[17] Olivero, Maria and Yoto Yotov (2011) "Dynamic Gravity: Theory and Empirical Implications," *Canadian Journal of Economics* forthcoming.

[18] Rust, John (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5).

[19] Santos Silva, Joao, and Silvana Tenreyro (2006). "The log of gravity," *Review of Economics and Statistics*, 88(4).

Table 1: Regression Results (Basic Specification)

Moving Cost

	Estim	***SE***
Mean C_t/ν	4.671 **	*(0.055)*
Max C_t/ν	4.884 **	*(0.059)*
Min C_t/ν	4.488 **	*(0.054)*
$1/\nu$	0.959 **	*(0.255)*

η/ν (Utility)

Sector	***Estim***	***SE***
1	0.000	-
2	-0.595 **	*(0.162)*
3	-0.156 *	*(0.094)*
4	-0.231 *	*(0.125)*
5	-0.224 *	*(0.114)*
6	-0.258 **	*(0.113)*
7	-0.601 **	*(0.166)*
8	-0.420 **	*(0.131)*
9	-0.297 **	*(0.122)*
10	-0.063	*(0.070)*
11	-0.486 **	*(0.169)*
12	-0.224 *	*(0.101)*
13	-0.140 **	*(0.052)*
14	-0.385 **	*(0.080)*
15	-0.285 *	*(0.133)*
16	-0.295 **	*(0.128)*

* significant at 5% level.
** significant at 1% level.

Table 2: Regression Results (Alternative Specification)

Moving Cost

	Estim	***SE***
Mean C_t/ν	4.671 **	*(0.055)*
Max C_t/ν	4.884 **	*(0.059)*
Min C_t/ν	4.488 **	*(0.054)*
$1/\nu$	3.672 **	*(0.667)*

η/ν (Utility)

	Intercept		Trend	
Sector	***Estim***	***SE***	***Estim***	***SE***
1	0.000	-	0.000	-
2	-2.378 **	*(0.455)*	0.009	*(0.008)*
3	-1.422 **	*(0.314)*	0.027 **	*(0.009)*
4	-1.528 **	*(0.340)*	0.003	*(0.008)*
5	-1.509 **	*(0.330)*	0.011	*(0.008)*
6	-1.690 **	*(0.359)*	0.023 **	*(0.009)*
7	-2.257 **	*(0.425)*	-0.004	*(0.008)*
8	-1.565 **	*(0.313)*	-0.014 *	*(0.008)*
9	-1.712 **	*(0.362)*	0.014 *	*(0.008)*
10	-0.857 **	*(0.224)*	0.014 *	*(0.008)*
11	-2.057 **	*(0.407)*	-0.013	*(0.008)*
12	-1.234 **	*(0.269)*	0.001	*(0.007)*
13	-0.727 **	*(0.166)*	0.017 *	*(0.008)*
14	-1.351 **	*(0.239)*	0.018 **	*(0.008)*
15	-1.497 **	*(0.329)*	-0.010	*(0.008)*
16	-1.505 **	*(0.331)*	-0.006	*(0.007)*

* significant at 5% level.
** significant at 1% level.

Table 3: Descriptive Statistics and Simulation Parameters

	Descriptive Statistics				Simulation Parameters		
	Labor Allocation		**Wage**		**Labor Share**	**Constant**	**CPI Share**
Sector	***Mean***	***SE***	***Mean***	***SE***	a	$\tilde{A}$	b
1	0.02	*(0.00)*	0.58	*(0.03)*	0.30	0.14	0.07
2	0.02	*(0.00)*	1.19	*(0.05)*	0.30	0.23	0.00
3	0.09	*(0.01)*	0.92	*(0.06)*	0.85	0.75	0.30
4	0.17	*(0.02)*	1.04	*(0.03)*	0.57	0.86	0.20
5	0.10	*(0.01)*	1.00	*(0.03)*	0.57	0.64	0.10
6	0.06	*(0.00)*	1.00	*(0.05)*	0.49	0.50	0.06
7	0.02	*(0.00)*	1.21	*(0.06)*	0.42	0.28	0.03
8	0.03	*(0.00)*	1.07	*(0.05)*	0.49	0.34	0.04
9	0.06	*(0.00)*	1.04	*(0.04)*	0.58	0.53	0.00
10	0.11	*(0.01)*	0.81	*(0.05)*	0.58	0.54	0.00
11	0.05	*(0.00)*	1.22	*(0.09)*	0.22	0.54	0.01
12	0.05	*(0.01)*	0.95	*(0.05)*	0.68	0.53	0.05
13	0.01	*(0.00)*	0.71	*(0.04)*	0.61	0.22	0.03
14	0.01	*(0.00)*	0.85	*(0.05)*	0.60	0.22	0.06
15	0.14	*(0.01)*	1.08	*(0.05)*	0.68	0.84	0.06
16	0.07	*(0.01)*	1.06	*(0.03)*	0.82	0.81	0.00

Table 4: Simulation Results: Estimation with PPML

Moving Cost (C/ν **and** $1/\nu$)									
		Sim I		**Sim II**		**Sim III**		**Sim IV**	
	Actual	***Estim***	***SE***	***Estim***	***SE***	***Estim***	***SE***	***Estim***	***SE***
Mean C_t/ν	4.500	4.503	*(0.023)*	4.503	*(0.022)*	4.503	*(0.022)*	4.504	*(0.037)*
Max C_t/ν	4.500	4.507	*(0.024)*	4.507	*(0.022)*	4.507	*(0.022)*	4.509	*(0.035)*
Min C_t/ν	4.500	4.500	*(0.021)*	4.497	*(0.021)*	4.497	*(0.021)*	4.498	*(0.036)*
$1/\nu$	1.000	0.995	*(0.119)*	0.993	*(0.109)*	0.999	*(0.049)*	1.010	*(0.186)*

Fixed Utility (η^i/ν)									
		Sim I		**Sim II**		**Sim III**		**Sim IV**	
Sector	***Actual***	***Estim***	***SE***	***Estim***	***SE***	***Estim***	***SE***	***Estim***	***SE***
2	0.100	0.103	*(0.017)*	0.101	*(0.016)*	0.101	*(0.007)*	0.101	*(0.020)*
3	0.150	0.153	*(0.020)*	0.153	*(0.017)*	0.151	*(0.008)*	0.151	*(0.024)*
4	0.200	0.202	*(0.019)*	0.203	*(0.017)*	0.201	*(0.007)*	0.201	*(0.025)*
5	0.250	0.252	*(0.016)*	0.252	*(0.016)*	0.250	*(0.007)*	0.252	*(0.019)*
6	0.300	0.301	*(0.020)*	0.301	*(0.017)*	0.301	*(0.008)*	0.303	*(0.023)*
7	0.350	0.352	*(0.027)*	0.350	*(0.026)*	0.351	*(0.011)*	0.352	*(0.039)*
8	0.400	0.400	*(0.029)*	0.399	*(0.027)*	0.400	*(0.012)*	0.403	*(0.040)*
9	0.000	0.002	*(0.027)*	0.002	*(0.025)*	0.000	*(0.013)*	-	-
10	-0.100	-0.098	*(0.036)*	-0.097	*(0.034)*	-0.100	*(0.015)*	-	-
11	-0.150	-0.148	*(0.040)*	-0.147	*(0.039)*	-0.149	*(0.018)*	-	-
12	-0.200	-0.197	*(0.040)*	-0.196	*(0.038)*	-0.199	*(0.017)*	-	-
13	-0.250	-0.248	*(0.022)*	-0.249	*(0.022)*	-0.251	*(0.009)*	-	-
14	-0.300	-0.298	*(0.024)*	-0.301	*(0.023)*	-0.302	*(0.012)*	-	-
15	-0.350	-0.346	*(0.064)*	-0.345	*(0.064)*	-0.350	*(0.029)*	-	-
16	-0.400	-0.396	*(0.059)*	-0.395	*(0.055)*	-0.399	*(0.025)*	-	-

Table 5: Simulation Results: Imputing Values with PPML and CCP

		$L = 4,000$		$L = 20,000$		$L \to \infty$	
	Actual	PPML	CCP	PPML	CCP	PPML	CCP
Λ^2	0.395	0.391	0.578	0.400	0.421	0.395	0.395
	(0.142)	*(0.305)*	*(0.327)*	*(0.185)*	*(0.244)*	*(0.142)*	*(0.142)*
Λ^3	1.025	1.030	1.378	1.026	1.020	1.025	1.025
	(0.145)	*(0.298)*	*(0.345)*	*(0.178)*	*(0.230)*	*(0.145)*	*(0.145)*
Λ^4	1.518	1.511	1.919	1.519	1.501	1.518	1.518
	(0.179)	*(0.295)*	*(0.339)*	*(0.198)*	*(0.249)*	*(0.179)*	*(0.179)*
Λ^5	1.248	1.262	1.642	1.251	1.252	1.248	1.248
	(0.138)	*(0.275)*	*(0.323)*	*(0.166)*	*(0.219)*	*(0.138)*	*(0.138)*
Λ^6	1.106	1.114	1.473	1.112	1.122	1.106	1.106
	(0.168)	*(0.276)*	*(0.313)*	*(0.194)*	*(0.235)*	*(0.168)*	*(0.168)*
Λ^7	0.716	0.730	1.009	0.713	0.730	0.716	0.716
	(0.153)	*(0.295)*	*(0.318)*	*(0.173)*	*(0.224)*	*(0.153)*	*(0.153)*
Λ^8	0.895	0.908	1.218	0.903	0.908	0.895	0.895
	(0.153)	*(0.271)*	*(0.298)*	*(0.182)*	*(0.234)*	*(0.153)*	*(0.153)*
Λ^9	0.797	0.823	1.143	0.785	0.794	0.797	0.797
	(0.155)	*(0.273)*	*(0.322)*	*(0.176)*	*(0.223)*	*(0.155)*	*(0.155)*
Λ^{10}	0.703	0.711	0.987	0.702	0.707	0.703	0.703
	(0.141)	*(0.282)*	*(0.334)*	*(0.176)*	*(0.229)*	*(0.141)*	*(0.141)*
Λ^{11}	0.736	0.745	1.028	0.738	0.757	0.736	0.736
	(0.146)	*(0.285)*	*(0.343)*	*(0.179)*	*(0.242)*	*(0.146)*	*(0.146)*
Λ^{12}	0.368	0.395	0.562	0.378	0.400	0.368	0.368
	(0.136)	*(0.270)*	*(0.319)*	*(0.183)*	*(0.249)*	*(0.136)*	*(0.136)*
Λ^{13}	-0.405	-0.416	-0.651	-0.407	-0.431	-0.405	-0.405
	(0.127)	*(0.344)*	*(0.447)*	*(0.174)*	*(0.259)*	*(0.127)*	*(0.127)*
Λ^{14}	-0.433	-0.452	-0.685	-0.425	-0.424	-0.433	-0.433
	(0.122)	*(0.343)*	*(0.468)*	*(0.190)*	*(0.252)*	*(0.122)*	*(0.122)*
Λ^{15}	0.840	0.837	1.137	0.840	0.846	0.840	0.840
	(0.156)	*(0.292)*	*(0.341)*	*(0.182)*	*(0.245)*	*(0.156)*	*(0.156)*
Λ^{16}	0.273	0.266	0.390	0.281	0.285	0.273	0.273
	(0.147)	*(0.301)*	*(0.363)*	*(0.194)*	*(0.242)*	*(0.147)*	*(0.147)*

Table 6: Simulation Results: Comparing Different Methods (with Aggregate Shocks)

Sample Size	Method	C/ν		$1/\nu$	
-	Actual	4.500	-	1.000	-
$L = 2,000$, $T = 25$	PPML1	4.530	*(0.015)*	1.001	*(0.024)*
	PPML2	4.515	*(0.016)*	1.006	*(0.027)*
	ACM	4.217	*(0.250)*	0.908	*(0.084)*
$L = 4,000$, $T = 25$	PPML1	4.515	*(0.010)*	1.000	*(0.020)*
	PPML2	4.506	*(0.010)*	1.003	*(0.021)*
	ACM	4.429	*(0.179)*	0.958	*(0.060)*
	CCP	4.517	*(0.011)*	1.083	*(0.038)*
$L = 20,000$, $T = 25$	PPML1	4.503	*(0.005)*	0.999	*(0.014)*
	PPML2	4.500	*(0.005)*	1.003	*(0.016)*
	ACM	4.560	*(0.074)*	1.001	*(0.032)*
	CCP	4.506	*(0.005)*	0.998	*(0.018)*
$L \to \infty$, $T = 25$	PPML1	4.500	*(0.000)*	0.999	*(0.012)*
	PPML2	4.498	*(0.001)*	1.003	*(0.014)*
	ACM	4.500	*(0.000)*	0.999	*(0.019)*
	CCP	4.498	*(0.001)*	1.003	*(0.014)*

Table 7: Simulation Results: Comparing Different Methods (without Aggregate Shocks)

Sample Size	Method	C/ν		$1/\nu$	
-	Actual	4.500	-	1.000	-
$L = 2,000$, $T = 25$	PPML1	4.530	*(0.015)*	0.999	*(0.023)*
	PPML2	4.515	*(0.015)*	1.007	*(0.023)*
	ACM	4.248	*(0.269)*	0.912	*(0.081)*
	PPML-NPM	4.501	*(0.014)*	1.013	*(0.016)*
$L = 4,000$, $T = 25$	PPML1	4.515	*(0.010)*	0.999	*(0.016)*
	PPML2	4.507	*(0.010)*	1.003	*(0.016)*
	ACM	4.429	*(0.177)*	0.961	*(0.056)*
	CCP	4.520	*(0.011)*	1.080	*(0.031)*
	NPM	4.495	*(0.010)*	1.006	*(0.015)*
$L = 20,000$, $T = 25$	PPML1	4.503	*(0.005)*	1.000	*(0.007)*
	PPML2	4.500	*(0.005)*	1.001	*(0.007)*
	ACM	4.559	*(0.073)*	1.006	*(0.025)*
	CCP	4.505	*(0.005)*	0.996	*(0.012)*
	NPM	4.499	*(0.005)*	1.015	*(0.016)*
$L \to \infty$, $T = 25$	PPML1	4.500	*(0.000)*	1.000	*(0.000)*
	PPML2	4.500	*(0.000)*	1.000	*(0.000)*
	ACM	4.500	*(0.000)*	1.000	*(0.000)*
	CCP	4.500	*(0.000)*	1.000	*(0.000)*
	NPM	4.500	*(0.000)*	1.000	*(0.000)*