A Unified Method for Dynamic and Cross-Sectional Heterogeneity: Introducing Hidden Markov Panel Models*

Jong Hee Park
Department of Political Science
University of Chicago

First Version March 25, 2009
This Version November 4, 2011

Abstract

Conventional statistical methods for panel data are based on the assumption that unobserved heterogeneity is time-constant. Despite the central importance of this assumption for panel data methods, few studies have developed statistical methods for testing this assumption and modeling time-varying unobserved heterogeneity. In this paper, I introduce a formal test to check the assumption of time-constant unobserved heterogeneity using Bayesian model comparison. Then, I present two panel data methods that account for time-varying unobserved heterogeneity in the context of the random-effects model and the fixed-effects model, respectively. I illustrate the utility of the introduced methods using both simulated data and examples drawn from two important debates in the political economy literature: (1) the identification of shifting relationships between income inequality and economic development in capitalist countries and (2) the effects of the GATT/WTO on bilateral trade volumes.

*`R` functions for all the proposed methods in this paper are provided within `MCMCpack` v.1.2-1 (Martin, Quinn and Park, 2011). Supplementary material of the paper is available at the author's website.

Introduction

In the social sciences, discontinuous changes, often identified as "periods," "critical junctures," "structural changes," or "turning points," play an important role in theory development and theory testing (e.g. Lucas, 1976; Gellner, 1992; Tilly, 1995; Abbott, 2001; Pierson, 2004; Sewell Jr., 2005). Whether these changes are a substance from which researchers derive important implications or a nuisance that impedes scientific inference, it is essential to identify changes in social processes when they exist.

The existence of hidden changes is a particularly alarming issue in panel data analysis, as conventional statistical methods for panel data are based on the assumption of time-constant unobserved heterogeneity.[1] Failing to account for changes in unobserved heterogeneity, when they exist, leads to omitted variable bias and hence results in invalid inferences. Specifically, when individual effects are time-varying, fixed-effects models and random-effects models produce inconsistent estimates.[2]

In this paper, I introduce a suite of statistical methods to test and model time-varying unobserved heterogeneity in the unified framework of a hidden Markov model (HMM). HMM is a statistical model that assumes a data-generating process governed by hidden regime transitions. It has been successfully applied to modeling heterogeneity in time series data, spatial data, and genetic data (Cappe, Moulines and Ryden, 2005). HMM has a strong theoretical appeal to social scientists as changes in social history are often identified as transitions between discrete regimes (Katznelson, 1997; Abbott, 2001; Lieberman, 2001;

[1]Another important assumption in conventional panel data methods is that unobserved individual effects are *additive*. I do not discuss this issue in this paper. See Bai (2009) for a method that addresses the interactive fixed-effects.

[2]See the appendix for the proof.

Sewell Jr., 2005; Western and Kleykamp, 2004; Wawro and Katznelson, 2011). Conventional statistical models assuming a stationary data generating process cannot address the full impacts of discontinuous social changes. We need a different tool to discover the existence and assess the impacts of these structural changes in the study of historical panel data.

First, I introduce two formal methods of testing the assumption of time-constant unobserved heterogeneity in conventional panel models. Residuals from conventional panel models are analyzed to check for the existence of hidden breaks at the group-level or at the subject-specific level using HMM.[3]

Then, I provide two statistical models to model the time-varying unobserved heterogeneity. The first model is the fixed-effects HMM. The fixed-effects HMM treats hidden breaks as a nuisance that hinders the consistent estimation of regression coefficients. In the fixed-effects HMM, unobserved factors are decomposed into *subject-specific and regime-constant* factors. The second model is the random-effects HMM in which all or some parts of random-effects model parameters vary over time at the group-level. While the number, timing, and nature of breaks vary across subjects in the fixed-effects HMM, the random-effects HMM assumes system-wide contemporaneous regime changes.

To check the utility of the proposed methods, I present three examples. The first example is from simulated data, and the next two examples are drawn from two important debates in the political economy literature: shifting relationships between income inequality and economic development across countries and over time and the effect of GATT/WTO on bilateral trade.

[3]In this paper, I use "subject" or "individual" to denote the cross-sectional unit in panel data and "group" to denote the collection of cross-sectional units.

A Unified Framework for Time-Varying Unobserved Heterogeneity

In this section, I present a *unified* framework to test and model time-varying heterogeneity. The approach proposed in this paper has a unified framework in the sense that all the statistical methods in this paper are based on the common framework of Bayesian HMM analysis. Also, my two statistical models for time-varying heterogeneity, corresponding to fixed-effects models and random-effects models, share the same HMM framework.

HMM is a well-known statistical approach to modeling level shifts or variance changes in a time series process.[4] The essential part of HMM is the assumption of conditional independence: the distribution of data, conditional upon hidden states, is independent. Thus, once I know the hidden states, the statistical inference for regression parameters is the same as conventional regression models. Then, the essential question in HMM analysis is how to identify the number, timing and type of hidden regimes.

Among many methods for the hidden state inference, Chib (1998)'s non-ergodic HMM (or multiple change-point model) has several advantages for social science data.[5] First, history in the social sciences is generally considered as non-recurrent processes full of discontinuities

[4]For useful references in the HMM literature, see MacDonald and Zucchini (1997), Cappe, Moulines and Ryden (2005), and Frühwirth-Schnatter (2006).

[5]The non-erdogic HMM is defined by its constraint on the uni-directional movement of hidden regimes. Let $p_{ij} = \Pr(s_t = j | s_{t-1} = i)$ be the probability of moving to state j from state i at time t when the state at $t-1$ is i. Then, Chib's non-ergodic transition matrix with M hidden regimes is

$$\mathbf{P}_{M \times M} = \begin{pmatrix} p_{11} & p_{12} & 0 & \dots & 0 \\ 0 & p_{22} & p_{23} & \dots & 0 \\ & & \dots\dots\dots\dots\dots\dots & & \\ 0 & 0 & 0 & p_{M-1,M-1} & p_{M-1,M} \\ 0 & 0 & & 0 & 1 \end{pmatrix}.$$

This transition matrix allows only a forward movement of hidden regimes.

(Abbott, 2001; Lieberman, 2001; Sewell Jr., 2005). Thus, the non-ergodic assumption of regime transitions in Chib (1998)'s method is consistent with the general understanding of regime changes in social history. Second, the non-ergodic transition matrix makes it easy to identify *multiple* hidden states by simplifying the initial probability of hidden states and by labeling regime-specific parameters in a non-reversible way. Last, the non-ergodic transition matrix can be used to model both recurrent and non-recurrent regime transitions as no constraint is imposed on parameters of non-adjacent states (e.g. state 1 and state 3). Thus, a regime switching process can be identified by comparing parameters of non-adjacent states after model fitting.

Tests of Unobserved Heterogeneity

In this section I present two tests to diagnose changes in unobserved factors (or individual effects) in linear panel data models: a test for group-level breaks and a test for subject-specific breaks. The purpose of the tests is to check for the existence of changes in unobserved factors from panel residuals. If unobserved factors are time-constant, residuals should not show any systematic change over time and across subjects.

The statistical technique for the tests is Bayesian model comparison using marginal likelihoods. In Bayesian model comparison, the marginal likelihood of model k, $p(\mathbf{y}|k)$, corresponds to the probability of observing the data ($\mathbf{y}$) given the model under consideration (k). The posterior probability of break number k can be computed as a ratio of the sum of all candidate HMMs' marginal likelihoods, holding a parametric model $\mathcal{M}$ constant. Let K be an upper threshold for break numbers. Then, the posterior probability of break number k is

$$p(\mathcal{M}_k|\mathbf{y}) = \frac{p(\mathbf{y}|\mathcal{M}_k)p(\mathcal{M}_k)}{\sum_{i=0}^{K} p(\mathbf{y}|\mathcal{M}_i)}. \tag{1}$$

The result of Bayesian model comparison using marginal likelihoods reports the relative plausibility of each model (or hypothesis) as a probability. Thus, no arbitrary threshold is necessary for statistical decisions. Also, Bayesian model comparison allows us to rid ourselves of the assumption of asymptotic normal error and the distinction between nested and non-nested model testing.[6]

Among many computational approaches that have been proposed to compute marginal likelihoods (Geweke, 1989; Newton and Raftery, 1994; Meng and Wong, 1996; Kass and Raftery, 1995; Chib, 1995; Chib and Jeliazkov, 2001; Green, 1995), I use Chib (1995)'s algorithm to compute marginal likelihoods from Gibbs sampling outputs because of its computational advantages.[7]

In the following, I present a test for group-level breaks using multivariate residuals at the group level and a test for subject-specific breaks using univariate residuals at the subject level.

Group-Level Residual Break Test

1. Fit a linear Gaussian multivariate HMM with M number of breaks using N (the number

[6]In contrast, most model comparison statistics, such as likelihood ratio tests, AIC or BIC are based on the asymptotic normality of the error, which does not hold in HMMs. See Frühwirth-Schnatter (2006, Ch. 4 and 5) for the discussion on the limitations of classical model selection tools in mixture models, a variant of which is HMM.

[7]Estimation details are discussed in the supplementary material due to space limitations.

of subjects) $\times 1$ vector of panel residuals $\hat{\mathbf{e}}_t$:

$$\hat{\mathbf{e}}_t \sim \begin{cases} \mathcal{N}(\boldsymbol{\alpha}_1, \sigma_1^2 \mathbf{I}_{T_i \times 1}), & \text{if } s_t = 1, \\ \vdots & \vdots \\ \mathcal{N}(\boldsymbol{\alpha}_M, \sigma_M^2 \mathbf{I}_{T_i \times 1}), & \text{if } s_t = M. \end{cases}$$

2. Compute the marginal likelihood of the M-component HMM.

3. Repeat all steps by varying M from 0 to an upper bound of break number (K).

4. Compute the posterior model probability for $\mathcal{M}_k$, $k = 0, \ldots, K$

Similarly, subject-specific breaks can be detected by applying a linear Gaussian HMM to subject-specific residuals from panel data models.

Subject-Level Residual Break Test

1. Fit a linear Gaussian univariate HMM with M_i number of breaks using a residual of a panel data model for subject i $\hat{e}_{it}$:

$$\hat{e}_{it} \sim \begin{cases} \mathcal{N}(\alpha_{i,1}, \sigma_{i,1}^2), & \text{if } s_{it} = 1, \\ \vdots & \vdots \\ \mathcal{N}(\alpha_{i,M_i}, \sigma_{i,M_i}^2), & \text{if } s_{it} = M_i. \end{cases}$$

2. Compute the marginal likelihood of the M_i-component HMM.

3. Repeat the above step by varying M_i from 0 to an upper bound of break number (K_i).

4. Compute the posterior model probability for $\mathcal{M}_k$, $k = 0, \ldots, K_i$.

5. Repeat all steps for $i = 1, \ldots, N$

As researchers do not have clear knowledge about whether regime changes are experienced in common or are different for each individual subject, it is important to visually inspect temporal patterns of panel residuals before tests.

Modeling Time-Varying Unobserved Heterogeneity

When hidden breaks exist in panel residuals, researchers need to respecify their statistical models to account for these hidden changes. In this section, I introduce two panel HMMs focusing on two canonical panel data methods: the random-effects model and the fixed-effects model. A general form for a panel data model can be written following Laird and Ware (1982)'s notation in which y_{it} denote an observation for subject i at t, $\mathbf{x}_{it}$ be the $k \times 1$ vector of regressors, $\mathbf{w}_{it}$ be the $q \times 1$ vector of the random-effects regressors, which is a subset of $\mathbf{x}_{it}$, and $\mathbf{b}_i$ denote the $q \times 1$ random-effects coefficient vector with variance-covariance matrix $\mathbf{D}$:

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{w}'_{it}\mathbf{b}_i + \varepsilon_{it}, \quad \varepsilon_{it} \sim \mathcal{N}(0, \sigma^2), \quad \mathbf{b}_i \sim \mathcal{N}(0, \mathbf{D}). \tag{2}$$

The above model can be modified to denote the fixed-effects model by setting $\mathbf{x}'_{it}\boldsymbol{\beta} = \alpha_i + \tilde{\mathbf{x}}'_{it}\tilde{\boldsymbol{\beta}}$ and $\mathbf{w}'_{it} = \mathbf{0}$ where α_i denotes the unobserved time-constant individual effect for subject i, $\tilde{\mathbf{x}}_{it}$ is a model matrix without the constant, and $\tilde{\boldsymbol{\beta}}$ is the parameter vector minus an intercept. I use the multivariate Gaussian distribution for the prior distribution of $\boldsymbol{\beta}$, the inverse gamma distribution for the prior distribution of σ^2, and the inverse wishart distribution for the prior distribution of $\mathbf{D}$.

The first panel HMM I propose is the random-effects HMM in which all subjects in the panel data are assumed to be exposed to a similar type of unknown system-level shocks. The model can be written as a mixture of M random-effects panel models sequentially applied to M time blocks as follows:

$$y_{it} = \begin{cases} \mathbf{x}'_{it}\boldsymbol{\beta}_1 + \mathbf{w}'_{it}\mathbf{b}_i + \varepsilon_{it}, & \mathbf{b}_i \sim \mathcal{N}(0, \mathbf{D}_1), \quad \varepsilon_{it} \sim \mathcal{N}(0, \sigma_1^2) \quad \text{for} \quad t_0 \leq t < \tau_1 \\ \vdots \quad \vdots & \vdots \quad \vdots \quad \vdots \\ \mathbf{x}'_{it}\boldsymbol{\beta}_M + \mathbf{w}'_{it}\mathbf{b}_i + \varepsilon_{it}, & \mathbf{b}_i \sim \mathcal{N}(0, \mathbf{D}_M), \quad \varepsilon_{it} \sim \mathcal{N}(0, \sigma_M^2) \quad \text{for} \quad \tau_{M-1} \leq t < T \end{cases} \tag{3}$$

where τ_i is the break point between regime $i-1$ and regime i.[8]

In the random-effects HMM, regime changes in unobserved factors are considered as systematic features of data, not as a noise, that should be included in statistical inference. Thus, all regression parameters are *interacted* with changes in hidden regimes. In this sense, the random-effects HMM is is similar to the regression model in which all predictors are interacted with dummy variables of hidden regimes. However, two major differences still exist between them. One is that unlike the dummy variable method the random-effects HMM does not require the prior knowledge of the number and timing of hidden regimes. Another difference is that the random-effects HMM allows *effects* of regime transitions to

[8]The random-effects HMM in this paper is different from other longitudinal HMMs such as Frühwirth-Schnatter and Kaufmann (2008) and Scott, James and Sugar (2005) by the use of a non-ergodic transition matrix. As I have mentioned above, the use of a non-ergodic transition matrix helps us avoid the label-switching problem and the over-fitting problem. The over-fitting problem refers to situations in which model comparison tools fail to detect redundant hidden states, which often results when hidden states are poorly identified (Frühwirth-Schnatter, 2006, Ch.4.).

vary across subjects, which will be estimated by $\mathbf{b}_i$ and $\mathbf{D}_m$.[9]

The second panel HMM is the fixed-effects HMM in which subjects are assumed to be exposed subject-specific shocks. The model can be succinctly written in matrix form as follows:

$$\mathbf{y}_i^* = \boldsymbol{\alpha}_{i,\cdot} + \tilde{\mathbf{X}}_i^* \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i, \quad \boldsymbol{\varepsilon}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_i) \tag{4}$$

$$\boldsymbol{\alpha}_{i,\cdot} = \underbrace{\begin{pmatrix} \alpha_{i,1} \\ \vdots \\ \alpha_{i,M_i} \end{pmatrix}}_{(T_i \times 1)}, \quad \boldsymbol{\Omega}_i = \underbrace{\begin{pmatrix} \sigma_{i,1}^2 & 0 & \dots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & 0 & \dots & \sigma_{i,M_i}^2 \end{pmatrix}}_{(T_i \times T_i)}. \tag{5}$$

where T_i is the number of time series observations in subject i and M_i is the number of hidden regimes in individual effects of subject i. Note that the subject-level residual break test should be done to obtain M_i prior to the estimation of the fixed-effects HMM.

Unlike the random-effects HMM, the fixed-effects HMM considers regime changes in individual effects as a nuisance that inhibits statistical inference of slope parameters ($\boldsymbol{\beta}$). The constant fixed-effects model is a special case of the fixed-effects HMM where the number of a hidden regime is 1 for all i: that is, $\alpha_{i,s_t} = 0$ and $\sigma_{i,s_t}^2 = \sigma^2$ for all i and t. Only in that special (and potentially rare) case, the fixed-effects model provides a consistent estimate of $\boldsymbol{\beta}$.[10]

[9] The sampling algorithms of the two panel HMMs are discussed in the appendix.

[10] The fixed-effects HMM in this paper is distinguished from fixed-effects models with varying individual effects in the econometrics literature (Ahn, Lee and Schmidt, 2001; Pesaran, 2006; Bai, 2009) in following ways. First, in the fixed-effects HMM, there is no constraint at the global level on the number, timing and magnitude of regime transitions within individual groups. In contrast, Ahn, Lee and Schmidt (2001)'s method assumes that the timing of the breaks is identical across groups. Also, Pesaran (2006) and Bai (2009) assume that the dynamics of unobserved factors (e.g. the shape of changes over time) are common across groups, even as their effects may vary from group to group. Second, whereas Pesaran (2006) and Bai (2009) address smooth transitions of unobserved factors over time, the fixed-effects HMM is designed for modeling

Applications

Monte Carlo Simulations

Before applying my methods to real-world examples, I will evaluate the performance of the methods using simulated data. First, I generate multiple panel data sets with one break to test the performance of the random-effects HMM. The cross-sectional sample size (N) starts at 10 and increases by 10 to 40. The time series sample size (T) starts at 20 and increases by 20 to 100. I set the timing of breaks in the middle of the sample.[11]

Table 1 around here

I compare the RMSEs of the random-effects HMM with the RMSEs of the constant random-effects model: the results of this comparison are reported in Table 1.[12] Overall, the RMSEs from the constant random-effects models (top) are always larger than the RMSEs from the random-effects HMM regardless of the number of subjects (N) and the length of

discontinuous changes in subject-specific hidden regimes.

[11]Specifically, I use the following model;

$$\begin{aligned}
\mathbf{p}_i &= \mathbf{1} - \Phi\left(\frac{\textit{Time}_i - \textit{Break Point}_i}{\textit{Transition Pace}_i}\right) \\
\mathbf{y}_i &= \mathbf{p}_i(\mathbf{X}_i'\boldsymbol{\beta}_1 + \mathbf{u}_i) + (\mathbf{1} - \mathbf{p}_t)(\mathbf{X}_i'\boldsymbol{\beta}_2 + \mathbf{u}_i) \\
\mathbf{u}_i &\sim \mathcal{N}_{T_i}(\mathbf{0}, \Sigma_i) \\
\Sigma_i &= \begin{cases} \sigma_1^2\mathbf{I} + \mathbf{W}_i\mathbf{D}_1\mathbf{W}_i' & \text{if } m = 1 \\ \sigma_2^2\mathbf{I} + \mathbf{W}_i\mathbf{D}_2\mathbf{W}_i' & \text{if } m = 2 \end{cases}
\end{aligned}$$

where $\mathbf{p}_i$ denotes a $T_i \times 1$ vector of probabilities of a hidden state being 1 for subject i. $\mathbf{1} - \mathbf{p}_i$ a $T_i \times 1$ vector of regime 2 state probabilities for subject i. $\textit{Time}_i$ is a vector of time indicators, $\textit{Break Point}_i$ is the timing of true breaks, and $\textit{Transition Pace}_i$ is the speed of regime transition for subject i. I used $\textit{Transition Pace}_i = 1$ for the simulation study of the random-effects HMM.

[12]I use Park (2011)'s RMSE. Suppose that the true β has a break at τ and I have posterior draws of hidden states $\hat{s}_t$. Then, I can find parameters for each t at gth simulation step, $\beta_{\hat{s}_t}^{(g)}$. Using them, I computed the RMSE for β as follows:

$$RMSE_\beta = \sqrt{\frac{1}{G}\sum_{g=1}^{G}\left(\frac{1}{\tau}\sum_{t=1}^{\tau}(\beta_{\hat{s}_t}^{(g)} - \beta_t)^2\right)} \quad (6)$$

where β_t is the true β at t and G is the total number of simulations.

time series observation (T). The gaps in the performance become larger as I have more data over time and across subjects.

The data for the simulation study of the fixed-effects HMM are similarly generated from the above except for the fact that the number of subjects without regime changes (N_0 in Table 2) vary in the simulation. That is, I allow only a portion of panel subjects to be exposed to individual shocks and see how it affects the estimation of slope parameters. Unobserved factors are generated from a Normal distribution with a large variance. I fixed the cross-sectional sample size at 40 while varying the size of time series observations from 40 to 100.

Table 2 around here

Table 2 compares the performance of different fixed-effects models based on their RMSEs. The true value of the slope is 1 in the simulation. The top line represents the ideal case in the dummy variable regression to control for effects of time-varying individual effects: researchers use the correct knowledge of the number and timing of subject-specific regime transitions. I use this best estimate from the OLS with dummy variables as a benchmark for the performance of the fixed-effects HMMs.

The results in the middle of Table 2 are from the fixed-effects HMM using the result of the subject-level residual break test as an input for break numbers. In this analysis, the number and timings of breaks are considered as *unknown* quantities that need to be estimated from data. It is impressive to see that the fixed-effects HMM using the result of the subject-level residual break test outperforms the OLS with subject and regime dummy variables in terms of RMSE.

One interesting fact I found from the simulation study of the fixed-effects HMM is that the subject-level residual break test sometimes over-identifies subject-level break numbers. As long as regime changes are non-deterministic, identifying the *true* numbers of breaks in a nuisance parameter does not make much sense, but it is important to check whether the estimation of slopes is sensitive to the over-fitting problem. The bottom of Table 2 shows the results of the fixed-effects HMMs that uses the *true* break numbers as an input. The comparison of RMSE between the two fixed-effects HMMs in the middle and bottom indicates that the prior specification of true break numbers does not necessarily improve the precision of slope estimates in the fixed-effects HMM.[13]

Turning Points in Inequality and Economic Development

Nearly six decades ago, Kuznets (1955) observed that inequality increases and subsequently falls during the process of industrialization. Since then, many political economists have investigated the inverse U-shaped relationship between inequality and development (Aghion, Caroli and Garcia-Penalosa, 1999; Alberto and Rodrik, 1994; Persson and Tabellini, 1994; Perotti, 1996; Acemoglu and Robinson, 2000; Boix, 2003). For example, Li and Zou (1998) find that inequality has a positive effect for economic growth, while Alberto and Rodrik (1994), Persson and Tabellini (1994), and Perotti (1996) find the opposite effect. Using a cross-national data set, Barro (2000) finds weak support for the Kuznets hypothesis, but reports that a large portion of the data remains unexplained.

[13] Although the full investigation of the over-fitting problem in HMM is beyond the scope of this paper, I have two explanations for this result. First, the step in my method that compares model fits using marginal likelihoods guards against excessive over-fitting. Second, additional breaks tend to be found when regime transitions were slow or smooth in the simulation. In these cases, parametric differences are small between adjacent regimes, and hence their effects on the estimation of slopes are minimal.

One possible cause of the conflicting findings is *changing relationships* between income inequality and economic development across countries and over time. In fact, recent political economy models view the relationships between inequality and economic development as nonlinear over time and non-monolithic across space (Acemoglu and Robinson, 2000, 2002; Boix, 2003, 2009; Rogowski and MacRae, 2008).[14] I investigate this conjecture using the random-effects HMM.

To measure income inequality across countries and over time, I use Leigh (2007)'s adjusted top 1 percent income share, excluding capital gains. Many studies show that the income share of a nation's richest population, measured from tax return data, is a highly reliable cross-national and historical measure of inequality (Piketty and Saez, 2006; Atkinson and Piketty, 2007; Leigh, 2007; Scheve and Stasavage, 2009). I use log GDP per capita obtained from Maddison (2010) as a measure of economic development. I use a balanced panel data set tracking seven developed countries – U.S., Canada, the Netherlands, France, Japan, Australia, and New Zealand – between the years 1923 and 1998.[15]

For the tests of time-constant unobserved heterogeneity, I first describe a relationship between income inequality and economic development as an error correction process and then add random-effects at the country level to capture the unobserved cross-national heterogeneity, resulting in the random-effects error correction model (ECM).[16]

[14] For example, after reviewing recent quantitative studies of income inequality in the social sciences, Atkinson and Brandolini (2009) conclude that modern findings in the study of income inequality are often "plagued by discontinuities which can seriously affect regression results" (Atkinson and Brandolini, 2009, 381).

[15] The choice of sample countries is limited by the availability of data on top income shares over a period of at least 50 years. For computational convenience, I trim the data set to construct a balanced panel.

[16] I chose to investigate the effect of economic development on income inequality, not vice versa. However, my goal is not to make causal inference but to identify the time-varying correlation between two time series variables. The error correction model is equivalent to the distributed lag model and chosen for easy interpretation following Keele and DeBoef (2008).

Figure 1 and Figure 2 around here

Figure 1 shows country-specific residuals from the random-effects ECM. Visual inspection indicates that country-specific residuals seem to have breaks around the 1940s and 1990s in most countries except New Zealand. The break test at the group-level rejects the hypothesis that the residuals have no break at the group level. Also, country-specific break tests rejected the hypothesis that unobserved factors are time-constant within each country except for New Zealand.[17]

The timings of country-specific breaks in the country-specific residuals are displayed in Figure 2. All countries except for New Zealand have breaks around the 1940s and the U.S. is the only country with another break in the mid 1980s. Thus, it seems reasonable to fit a random-effects HMM with one break to investigate the shifting relationships between income inequality and economic development across countries. The resulting random-effects ECM with the hidden regime m at t is

$$\begin{aligned}\Delta\text{Inequality}_{it} &= \alpha_{i,m} + \beta_{1,m}\text{Inequality}_{it-1} + \beta_{2,m}\Delta\text{logGDP}_{it} + \beta_{3,m}\text{logGDP}_{it-1} + \varepsilon_{it}, \quad (7)\\ \varepsilon_{it} &\sim \mathcal{N}(0, \sigma^2_m), \quad \alpha_{i,m} \sim \mathcal{N}(\alpha_0, \sigma^2_{\alpha,m}).\end{aligned}$$

The Bayesian model comparison of random-effects ECMs from 0 to 4 breaks shows that the two break model is most reasonable with the posterior model probability of .81.[18] Table 3 compares parameter estimates of the two-break random-effects ECM with parameter estimates of the random-effects ECM with no break. If we compare estimates of the short-term

[17] The detailed test results are reported in the supplementary material.

[18] The test results are $\log p(\mathbf{y}|k = 0) = -528.52$, $\log p(\mathbf{y}|k = 1) = -500.89$, $\log p(\mathbf{y}|k = 2) = -499.27$, $\log p(\mathbf{y}|k = 3) = -502.66$, $\log p(\mathbf{y}|k = 4) = -528.40$.

effect of economic growth (*Growth*) on top 1 percent income share, the random-effects ECM reports that an additional percentage point of economic growth increases the top 1 percent income share by 1.80 percent during the 20th century. Thus, the substantive conclusion from the constant random-effects ECM would be that income growth always accompanies greater inequality in the economy in the history of 20th century capitalist development. To gauge the substantive effect of this estimate, consider the fact that the largest annual change in the U.S. top 1 percent income share in the 20th century was 2 percent between 1929 and 1930 (from 18.42 to 16.42). However, this pessimistic conclusion is unwarranted from the Bayesian viewpoint because the posterior probability of the zero-break model is effectively 0.

Table 3 around here

Countries have been exposed to various shocks during the 20th century and if we account for the unmodeled effects of these common shocks using HMM, we reach a very different conclusion as reported in the bottom three lines of Table 3. Most interestingly, the short-term effect of economic growth on top 1 percent income share shifts dramatically across three distinct regimes. While advanced economies during the first regime (between 1942 and 1952) and the third regime (between 1987 and 1998) experienced inegalitarian economic growth with large volatility, the second regime (between 1953 and 1986) can be characterized as a period of egalitarian growth. During the second regime, an additional percentage point of economic growth does not significantly increase the top 1 percent income share.

Another interesting time-varying pattern can be found from the residual variance (σ^2). Again, the second regime is characterized as the stable relationship between the top 1 percent income share and economic growth. The small slopes and the small residual variance during the second regime indicate that the top 1 percent income share stayed very close to the intercept during these periods. Overall, the tale of inequality and economic growth from the random-effects HMM is quite different from the one I inferred from the constant random effects model.

Effects of the GATT/WTO on Bilateral Trade

The next example comes from the debate on whether the multilateral trade organizations have significant effects on international trade.[19] Among these, I revisit a study by Goldstein, Rivers and Tomz (2007, hereafter GRT) that finds significant effects of the GATT/WTO on bilateral trade. Most of the studies involved in this debate – including GRT – use the fixed-effects method to control for dyad-specific and year-specific unobserved factors in bilateral trade data. I test the validity of the time-constant individual effect assumption and show how the results change when I relax the assumption of time-constant individual effects.

It is computationally challenging to directly apply the fixed-effects HMM to the entire data set of GRT with 17,359 dyads and 381,656 observations. Thus, I first show the results of the fixed-effects HMM within industrial dyads (594 dyads and 28,971 observations) and then employ the OLS with dyad and regime fixed-effects to the entire data set. As shown in the simulation study, results from the OLS with dyad and (correctly specified) regime fixed-

[19]Rose (2003) initiated the debate by publishing a paper that questions the effect of the GATT/WTO on trade volumes. Analyzing bilateral trade data over 175 countries and 50 years, Rose concluded that "the GATT/WTO seems to have a huge effect on trade if one does not hold other things constant; the multilateral trade regime matters, ceteris **non** paribus" (emphasis original, Rose, 2003, 111). This finding quickly generated subsequent studies, either modifying or criticizing the claim of a null effect (e.g. Gowa and Kim, 2005; Tomz, Goldstein and Rivers, 2007; Subramanian and Wei, 2007; Goldstein, Rivers and Tomz, 2007).

effects are close to the results from the fixed-effects HMM. Dyad-specific hidden regimes are identified by the subject-level residual break test for all dyads with long time series data.

The subject-level residual break test for industrial dyads finds that 500 dyads have one (154 dyads), two (283 dyads) or three breaks (63 dyads) in their residuals among 505 dyads with longer than 20 years of time series observations. Figure 3 reports differences in parameter estimates between GRT (empty circles) and the fixed-effects HMM (dark solid circles) in industrial dyads.

Figure 3 and Table 4 around here

Overall, the distances between the two estimates are substantial and some of the parameters move close to zero when I account for regime changes in individual effects using the fixed-effects HMM. To gauge the substantive effects of these movements, I report the effects of each covariate in Table 4. First, while GRT report that "the GATT/WTO expanded commerce by more than 70 percent when both trading partners were industrial nations," the fixed-effects HMM reports a much smaller effect of 10.6 percent. Similarly, the effect of GATT/WTO for "Only one participates in the GATT/WTO" shrinks by half when I control for country-specific hidden regime changes in individual effects through the fixed-effects HMM. Most interestingly, when I account for country-specific hidden regime changes in individual effects, GRT's conclusion that industrial countries gained more from participation in the GATT/WTO than from preferential trade agreements (PTAs) seems unwarranted. Instead, the results from the fixed-effects HMM suggest the opposite conclusion: PTAs benefits industrial countries more than the GATT/WTO.

In order to check whether the findings from industrial dyads are exceptional, I implemented the subject-level residual break test for 3,892 dyads with longer than 40 years of observations among all dyads. 2,541 dyads were detected to have one break and 1,137 dyads were detected to have two breaks. Using the results, I fit the OLS with dyad and regime fixed-effects to control for the dyad and regime-specific unobserved factors. The results are shown in Table 5.

Table 5 around here

As in the case of industrial dyads, the positive effects of the GATT/WTO participation on bilateral trade significantly decrease when dyad and regime-specific unobserved factors are removed from the analysis using the fixed-effects method. Parameter changes in other variables (e.g. Reciprocal PTA, Nonreciprocal PTA, GDP) are rather small as in the case of industrial dyads.[20]

Concluding Remarks

In this paper, I introduced a suite of statistical methods for diagnosing and modeling changes in unobserved heterogeneity. The foregoing simulation study showed that failing to account for changes in unobserved heterogeneity in panel data analysis produce highly inaccurate parameter estimates in terms of the mean squared error. The investigation of turning points

[20]The sensitivity to different dynamic specifications can also be found from GRT's time-varying analysis (Table 8 in GRT). GRT estimated the time-varying effects of covariates using a natural cubic spline function. In estimating coefficients of the cubic spline functions, they also included *year fixed effects* in the model. One may disagree about the use of year fixed-effects in estimating time-varying effects of covariates based on the argument that I cannot treat time both as both a substance and a nuisance simultaneously. When I dropped year dummy variables and re-estimated their natural cubic spline coefficients, the effects of the GATT/WTO participation variables decreased significantly. For example, the effect of "Both participate in the GATT/WTO" in the 1950s moves from 85 percent to 56 percent when year dummy variables are dropped from the model. The results are available in the supplementary material. Another potential source of the sensitivity is the natural cubic spline function, which requires users to pre-define the number and position of knots.

in the relationship between inequality and economic development showed that the random-effects HMM is a highly useful method for detecting and analyzing parametric changes due to group-level shocks. Meanwhile, the application of the fixed-effects HMM to studying the effects of multilateral trade organizations on bilateral trade illustrated that accounting for changes in individual effects may change the substantive conclusions of analyses based on the time-constant fixed-effects method.

The models introduced in this paper are neither complete nor all-encompassing. The discussion in this paper is limited to the linear Gaussian panel model with a simple nested panel structure. Given the complexity of data-generating processes of historical panel data in the social sciences, extensions to discrete response data and complex panel structure will be important topics for future research. The goal of this paper is to provide a unified framework for this future research.

References

Abbott, Andrew. 2001. *Time Matters: On Theory and Method*. University of Chicago Press.

Acemoglu, Daron, and James A. Robinson. 2002. "The Political Economy of the Kuznets Curve." *Review of Developmental Economics* 6:183–203.

Acemoglu, Daron, and James Robinson. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality and Growth in Historical Perspective." *Quarterly Journal of Economics* 115:1167–1199.

Aghion, Philippe, Eve Caroli, and Cecilia Garcia-Penalosa. 1999. "Inequality and Economic Growth: The Perspective of the New Growth Theories." *Journal of Economic Literature* 37:1615–1660.

Ahn, Seung C., Young H. Lee, and Peter Schmidt. 2001. "GMM Estimation of Linear Panel Data Models with Time-Varying Individual Effects." *Journal of Econometrics* 101:219–255.

Alberto, Alesina, and Dani Rodrik. 1994. "Distributive Politics and Economic Growth." *Quarterly Journal of Economics* 109:405 – 490.

Atkinson, Anthony Barnes, and A. Brandolini. 2009. "On Data: a case study of the evolution of income inequality across time and across countries." *Cambridge Journal of Economics* 33:381 – 404.

Atkinson, Anthony Barnes, and Thomas Piketty. 2007. *Top Incomes over the Twentieth Century: A Contrast between European and English-Speaking Countries*. Oxford University Press.

Bai, Jushan. 2009. "Panel Data Models With Interactive Fixed Effects." *Econometrica* 77:1229–1279.

Barro, Robert J. 2000. "Inequality and Growth in a Panel of Countries." *Journal of Economic Growth* 5 (March):5–32.

Boix, Carles. 2003. *Democracy and Redistribution*. Cambridge University Press.

Boix, Carles. 2009. "The Conditional Relationship between Inequality and Development." *PS: Political Science and Politics* 42:645–649.

Cappe, Oliver, Eric Moulines, and Tobias Ryden. 2005. *Inference in Hidden Markov Models*. Springer-Verlag.

Chib, Siddhartha. 1995. "Marginal Likelihood From the Gibbs Output." *Journal of the American Statistical Association* 90 (December):1313–1321.

Chib, Siddhartha. 1998. "Estimation and Comparison of Multiple Change-Point Models." *Journal of Econometrics* 86 (June):221–241.

Chib, Siddhartha, and Ivan Jeliazkov. 2001. "Marginal Likelihood from the Metropolis-Hastings Output." *Journal of the American Statistical Association* 96:270–281.

Frühwirth-Schnatter, Sylvia. 2006. *Finite Mixture and Markov Switching Models.* Springer-Verlag.

Frühwirth-Schnatter, Sylvia, and Sylvia Kaufmann. 2008. "Model-Based Clustering of Multiple Time Series." *Journal of Business & Economic Statistics* 26:78–89.

Gellner, Ernest. 1992. *Plough, Sword, and Book: The Structure of Human History.* University of Chicago Press.

Geweke, John. 1989. "Bayesian Inference in Econometric Models Using Monte Carlo Integration." *Econometrica* 57 (11):1317–1339.

Goldstein, Judith, Douglas Rivers, and Michael Tomz. 2007. "Institutions in International Relations: Understanding the Effects of the GATT and the WTO on World Trade." *International Organization* 61 (Winter):37–67.

Gowa, Joanne S., and Soo Yeon Kim. 2005. "An Exclusive Country Club: The Effects of the GATT on Trade, 1950-94." *World Politics* 57:453–478.

Green, Peter J. 1995. "Reversible Jum Markov Chain Monte Carlo Computation and Bayesian Model Determination." *Biometrika* 82:711–732.

Kass, Robert E., and Adrian E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90:773–795.

Katznelson, Ira. 1997. "Reflections on history, method, and political science." *Political Methodologist* 8:11–14.

Keele, Luke, and Suzanna DeBoef. 2008. "Taking Time Seriously: Dynamic Regression." *American Journal of Political Science* 52:184–200.

Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *American Economic Review* 45:1–28.

Laird, Nan M., and James H. Ware. 1982. "Random-Effects Models for Longitudinal Data." *Biometrics* 38:963–974.

Leigh, Andrew. 2007. "How Closely Do Top Income Shares Track Other Measures of Inequality?" *Economic Journal* 117:619–633.

Li, Hongyi, and Heng-fu Zou. 1998. "Income Inequality Is Not Harmful for Growth: Theory and Evidence." *Review of Developmental Economics* 2:318–334.

Lieberman, Evan S. 2001. "Causal Inference in Historical Institutional Analysis: A Specification of Periodization Strategies." *Comparative Political Studies* 34:1011–1035.

Lucas, Jr, Robert E. 1976. "Econometric policy evaluation: A critique." *Carnegie-Rochester Conference Series on Public Policy* 1:19–46.

MacDonald, Iain L., and Walter Zucchini. 1997. *Hidden Markov and Other Models for Discrete-valued Time Series*. Chapman and Hall.

Maddison, Angus. 2010. "Historical Statistics (`http://www.ggdc.net/maddison/`).".

Martin, Andrew D., Kevin M. Quinn, and Jong Hee Park. 2011. "`MCMCpack`, Version 1.1-4." `http://mcmcpack.wustl.edu/`.

Meng, Xiao-Li, and Wing Hung Wong. 1996. "Simulating Ratios of Normalizing Constants via a Simple Identity: A Theoretical Exploration." *Statistica Sinica* 6:831–860.

Newton, Michael A., and Adrian E. Raftery. 1994. "Approximate Bayesian Inference with the Weighted Likelihood Bootstrap." *Journal of the Royal Statistical Society. Series B (Methodological)* 56:3–48.

Park, Jong Hee. 2011. "Changepoint Analysis of Binary and Ordinal Probit Models: An Application to Bank Rate Policy Under the Interwar Gold Standard." *Political Analysis* 19:188 – 204.

Perotti, Roberto. 1996. "Growth, income distribution, and democracy: What the data say." *Journal of Economic Growth* 1:149–187.

Persson, Torsten, and Guido Tabellini. 1994. "Is Inequality Harmful for Growth?" *American Economic Review* 84:600–621.

Pesaran, M. Hashem. 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica* 74:967–1012.

Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. New Jersey: Princeton University Press.

Piketty, Thomas, and Emmanuel Saez. 2006. "The Evolution of Top Incomes: A Historical and International Perspective." *The American Economic Review* 96 (05):200–205.

Rogowski, Ronald, and Duncan C. MacRae. 2008. Inequality and Institutions: What Theory, History, and (some) Data Tell Us. In *Democracy, Inequality, and Representation*, ed. Pable Beramendi, and Christopher J. Anderson. Russel Sage.

Rose, Andrew. 2003. "Do We Really Know that the WTO Increases Trade?" *American Economic Review* 94:98–114.

Scheve, Kenneth, and David Stasavage. 2009. "Institutions, Partisanship, and Inequality in the Long Run." *World Politics* 61:215–253.

Scott, Steven L., Gareth M. James, and Catherine A. Sugar. 2005. "Hidden Markov Models for Longitudinal Comparisons." *Journal of the American Statistical Association* 100:359–369.

Sewell Jr., William H. 2005. *Logics of History: Social Theory and Social Transformation.* University of Chicago Press.

Subramanian, Arvind, and Shang-Jin Wei. 2007. "The WTO promotes trade, strongly but unevenly." *Journal of International Economics* 72:151 – 175.

Tilly, Charles. 1995. "To Explain Political Processes." *American Journal of Sociology* 100:1594 – 1610.

Tomz, Michael, Judith Goldstein, and Douglas Rivers. 2007. "Do We Really Know That the WTO Increases Trade? Comment." *American Economic Review* 97:2005 – 2018.

Wawro, Gregory J., and Ira Katznelson. 2011. "Political Science and History: Enhancing the Methodological Repertoire." Unpublished.

Western, Bruce, and Meredith Kleykamp. 2004. "A Bayesian Change Point Model for Historical Time Series Analysis." *Political Analysis* 12.

Software Implementation

```
require(MCMCpack)
## 1. subject specific break test
subject.break.test <- testpanelSubjectBreak(subject.id, time.id, residual)

## 2. group break test using zero break, one break and two break model
group.break0 <- testpanelGroupBreak(subject.id, time.id, residual, m=0)
group.break1 <- testpanelGroupBreak(subject.id, time.id, residual, m=1)
group.break2 <- testpanelGroupBreak(subject.id, time.id, residual, m=2)
print(BayesFactorList(list(group.break0, group.break1, group.break2)))

## 3. Random-effects HMMs with zero break, one break, and two breaks
RE0 <- HMMpanelRE(subject.id, time.id, y, X, W, m=0)
RE1 <- HMMpanelRE(subject.id, time.id, y, X, W, m=1)
RE2 <- HMMpanelRE(subject.id, time.id, y, X, W, m=2)
print(BayesFactorList(list(RE0, RE1, RE2)))

## 4. Fixed-effects HMM using group centered data and a break list
FE <- HMMpanelFE(subject.id, time.id, y=centered.y, X=centered.X, m=break.list)
```

Inconsistency of panel estimates in the presence of time-varying unobserved heterogeneity

Suppose that there were exogenous shocks at τ_1 that transform unobserved individual effects $\alpha_{i,t}$. Then, the true data-generating process is

$$y_{it} = \alpha_{i,t} + \mathbf{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it}, \;\; \varepsilon_{it} \sim \mathcal{N}(0, \sigma_y^2) \quad \text{For } i = 1, \ldots, N, t = 1, \ldots T. \tag{8}$$

where $\mathbf{x}_{it}$ indicates a vector of covariates observed from a subject i at time t.

$$E(\alpha_{i,t} - \bar{\alpha}_i) \;\; = \;\; \begin{cases} (1-\omega)(\alpha_{i,1} - \alpha_{i,2}) & \text{for } \; t_0 \leq t < \tau_1 \\ \omega(\alpha_{i,2} - \alpha_{i,1}) & \text{for } \; \tau_1 \leq t < T_i \end{cases}$$

and consequently,

$$\lim_{N \to \infty} \Pr \left(\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'(\alpha_{i,t} - \bar{\alpha}_i + \varepsilon_{it} - \bar{\varepsilon}_{it}) \right) \neq \mathbf{0}$$

To show the inconsistency of random-effects estimates, suppose that the exogenous shock at τ_1 transforms the distribution of individual effects in (8) as follows:

$$\alpha_{i,t} = a + \epsilon_{i,t}, \;\; \epsilon_{i,t} \;\; \sim \;\; \begin{cases} \mathcal{N}(0, \sigma_{\alpha_1}^2) & \text{for } \; t_0 \leq t < \tau_1 \\ \mathcal{N}(0, \sigma_{\alpha_2}^2) & \text{for } \; \tau_1 \leq t < T_i \end{cases}$$

Then, the variance of error terms is different before and after the break:

$$Var(\epsilon_{i,t} + \varepsilon_{it}) \;\; = \;\; \begin{cases} \sigma_{\alpha_1}^2 + \sigma_y^2 & \text{for } \; t_0 \leq t < \tau_1 \\ \sigma_{\alpha_2}^2 + \sigma_y^2 & \text{for } \; \tau_1 \leq t < T_i \end{cases}$$

Therefore, as long as $\sigma_{\alpha_1}^2 \neq \sigma_{\alpha_2}^2$, which is assumed to be true from the existence of the exogenous shock, the time-constant random effects estimate of $Var(\epsilon_{i,t} + \varepsilon_{it})$ is inconsistent, which leads to an inconsistent estimate of $\boldsymbol{\beta}$.

Algorithm 1. Random-Effects Hidden Markov Model

1. Sample $\mathbf{s}$ using Chib (1998)'s forward-backward recursions after rewriting the random-effects model into a multivariate time series model as follows:

$$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}_t\mathbf{b}_{[i]} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \sigma^2\mathbf{I}_{N_t}).$$

   N_t is the number of groups at t, $\mathbf{y}_t$ is the $N_t \times 1$ vector of the response variable at t, $\mathbf{X}_t$ is the $N_t \times K$ matrix, and $\mathbf{W}_t\mathbf{b}_{[i]}$ is $N_t \times 1$ matrix constructed by the subject-wise multiplication of $\mathbf{w}_{it}$ and $\mathbf{b}_i$.

2. Sample p_{kk} from $\mathcal{B}eta(a_0 + j_{k,k} - 1, b_0 + j_{k,k+1})$. p_{kk} is the probability of staying when the state is k, and $j_{k,k}$ is the number of jumps from state k to k, and $j_{k,k+1}$ is the number of jumps from state k to $k+1$.

3. Sample $\boldsymbol{\beta}_m$ from $\mathcal{N}(\hat{\boldsymbol{\beta}}_m, \mathbf{B}_m)$ where $\hat{\boldsymbol{\beta}}_m = \mathbf{B}_m(\mathbf{B}_0^{-1}\boldsymbol{\beta}_0 + \mathbf{X}'_{im}\mathbf{V}_{im}^{-1}\mathbf{y}_{im})$, $\mathbf{B}_m = (\mathbf{B}_0^{-1} + \mathbf{X}'_{im}\mathbf{V}_{im}^{-1}\mathbf{X}_{im})^{-1}$, and $\mathbf{V}_{im} = \mathbf{W}_{im}\mathbf{D}'_m\mathbf{W}_{im} + \sigma_m^2\mathbf{I}_{T_m}$. $\mathbf{X}_{im}$ and $\mathbf{y}_{im}$ are a $T_m \times K$ matrix and a $T_m \times 1$ matrix, respectively. T_m denotes the number of observations at state m for each group and K be the number of fixed-effects covariates including the intercept.

4. Sample $\mathbf{b}_{im}$ from $\mathcal{N}(\hat{\mathbf{b}}_{im}, \hat{\mathbf{D}}_m)$ where $\hat{\mathbf{b}}_{im} = \hat{\mathbf{D}}_m(\sigma_m^{-2}\mathbf{W}'_{im}(\mathbf{y}_{im} - \mathbf{X}_{im}\boldsymbol{\beta}_m))$ and $\hat{\mathbf{D}}_m = (\mathbf{D}_m^{-1} + \mathbf{W}'_{im}\mathbf{W}_{im}\sigma_m^{-2})^{-1}$.

5. Sample $\mathbf{D}_m^{-1}$ from $\mathcal{W}(\rho_0 + n_i, \mathbf{R}_m)$ where $\mathbf{R}_m = (\mathbf{R}_0^{-1} + \sum_{i=1}^{T_m}\mathbf{b}_{im}\mathbf{b}'_{im})^{-1}$ and n_i is the number of groups and T_m is the number of observations at state m as defined above.

6. Sample σ_m^2 from $\mathcal{IG}(\frac{v_0 + \sum_{i=1}^{n_i} T_m}{2}, \frac{v_1 + \hat{v}}{2})$ where $\hat{v} = \sum_{i=1}^{n_i}(\mathbf{y}_{im} - \mathbf{X}_{im}\boldsymbol{\beta}_m - \mathbf{W}_{im}\mathbf{b}_{im})'(\mathbf{y}_{im} - \mathbf{X}_{im}\boldsymbol{\beta}_m - \mathbf{W}_{im}\mathbf{b}_{im})$.

Algorithm 2. Fixed-Effects Hidden Markov Model

Before fitting the fixed-effects HMM, subject-specific break numbers M_i should be identified from the subject-level residual break test using residuals from the fixed-effect model.

1. Sample $\boldsymbol{\beta}$ from $\mathcal{N}(\hat{\boldsymbol{\beta}}, \mathbf{B})$ where $\mathbf{B} = (\mathbf{B}_0^{-1} + \sum_{i=1}^N \tilde{\mathbf{X}}_i^{*\prime}\tilde{\mathbf{X}}_i^*/\sigma^2)^{-1}$ and $\hat{\boldsymbol{\beta}} = \mathbf{B}(\mathbf{B}_0^{-1}\boldsymbol{\beta}_0 + \sum_{i=1}^N \tilde{\mathbf{X}}_i^{*\prime}\hat{\mathbf{y}}_i/\sigma^2)$. $\hat{\mathbf{y}}_i = \mathbf{y}_i^* - \boldsymbol{\alpha}_{i,\cdot}$ where $\boldsymbol{\alpha}_{i,\cdot}$ is a vector of time-varying intercepts for subject i. $\mathbf{y}_i^*$ is the centered response data for i and $\tilde{\mathbf{X}}_i^*$ are the centered model matrix without the constant term.

2. Sample $\mathbf{s}_i$ from $p(s_{it}|\check{\mathbf{y}}_i, \mathbf{S}_i^{t+1}, \mathbf{P}_i, \boldsymbol{\beta}, \sigma^2_{i,s_t}, M_i)$ where $\check{\mathbf{y}}_i = \mathbf{y}_i^* - \mathbf{X}_i'^*\boldsymbol{\beta}$.

3. Sample p_{ik} from $\mathcal{B}eta(j_0 + j_{k,k} - 1, j_1 + j_{k,k+1})$ where $j_{i,k}$ is the number of jumps from state k to k and $j_{i,k+1}$ is the number of jumps from state k to $k+1$.

4. Sample $\alpha_{i,m}$ from $\mathcal{N}(\hat{\alpha}_{i,m}, \mathbf{A}_{i,m})$ where $\hat{\alpha}_{i,m} = \mathbf{A}_{i,m}\left(\Delta^{-1}\delta + \sigma_{i,m}^{-2}\sum_{s_t=m}\tilde{y}_{it}\right)$ and $\mathbf{A}_{i,m} = \left(\Delta^{-1} + \sigma_{i,m}^{-2}N_{im}\right)^{-1}$ by looping m from 1 to $M_i + 1$. Note that $\tilde{y}_{it} = y_{it}^* - \mathbf{x}_{it}'^*\boldsymbol{\beta}$, N_{im} is a number of observations in state m, and $\sum_{s_t=m}$ is the summation over state m.

5. Sample $\sigma_{i,m}^2$ from $\mathcal{IG}(\frac{v_0 + N_{im}}{2}, \frac{v_1 + \hat{v}_{i,m}}{2})$ where $\hat{v}_{i,m} = (\tilde{\mathbf{y}}_{i,m} - \mathbf{X}_{i,m}^*\boldsymbol{\beta})'(\tilde{\mathbf{y}}_{i,m} - \mathbf{X}_{i,m}^*\boldsymbol{\beta})$. N_{im} is the number of observations at subject j's state m. $\mathbf{X}_{i,m}^*$ and $\tilde{\mathbf{y}}_{i,m}$ are a $N_{im} \times K$ matrix and a $N_{im} \times 1$ matrix, respectively.

Table 1: Root Mean Squared Error of the Constant Random-Effects Model and Random-Effects Hidden Markov Model from Simulated Data

Constant Random-Effects	RMSE				
	$T = 20$	$T = 40$	$T = 60$	$T = 80$	$T = 100$
$N = 10$	1.096	1.065	1.105	1.044	1.057
$N = 20$	1.104	1.064	1.055	1.026	1.031
$N = 30$	1.025	1.023	1.020	1.019	1.020
$N = 40$	1.020	1.026	1.015	1.017	1.014
Random-Effects HMM	RMSE				
	$T = 20$	$T = 40$	$T = 60$	$T = 80$	$T = 100$
$N = 10$	0.468	0.271	0.331	0.183	0.228
$N = 20$	0.340	0.263	0.347	0.195	0.240
$N = 30$	0.422	0.382	0.253	0.145	0.161
$N = 40$	0.368	0.220	0.321	0.165	0.123

Note: The MCMC iteration for each simulation is 1,000 after 1,000 burnin iterations. Prior distributions are $\boldsymbol{\beta} \sim \mathcal{N}(0, \sqrt{10}), \sigma^2 \sim \mathcal{IG}(1, 1), \mathbf{D} \sim \mathcal{W}(5, \hat{V}_{\text{OLS}}^{-1})$ where $\hat{V}_{\text{OLS}}^{-1}$ is the inverted variance-covariance matrix estimated from the OLS regression of $\mathbf{y}$ on $\mathbf{W}$. Prior distributions for transition probabilities are designed so that the expected regime duration evenly divides the sample period, reflecting my lack of knowledge about the timing of the break before I see data: $p_{ii} \sim \mathcal{B}eta(T/20, 0.1)$.

Table 2: Root Mean Squared Error of the Constant Fixed-Effects Model and Fixed-Effects Hidden Markov Model from Simulated Data

Model	Subject and Regime Fixed-Effects (OLS)			
Break Numbers	Known			
Break Timings	Known			
	$T = 40$	$T = 60$	$T = 80$	$T = 100$
$N_0 = 0$	0.249	0.246	0.199	0.160
$N_0 = 10$	0.192	0.177	0.175	0.130
$N_0 = 20$	0.268	0.310	0.154	0.209
$N_0 = 30$	0.130	0.142	0.061	0.110
Model	Fixed-Effects HMM			
Break Numbers	Unknown			
Break Timings	Unknown			
	$T = 40$	$T = 60$	$T = 80$	$T = 100$
$N_0 = 0$	0.067	0.037	0.025	0.030
$N_0 = 10$	0.037	0.025	0.027	0.023
$N_0 = 20$	0.032	0.023	0.026	0.017
$N_0 = 30$	0.068	0.023	0.018	0.017
Model	Fixed-Effects HMM			
Break Numbers	Known			
Break Timings	Unknown			
	$T = 40$	$T = 60$	$T = 80$	$T = 100$
$N_0 = 0$	0.067	0.047	0.027	0.031
$N_0 = 10$	0.039	0.025	0.028	0.024
$N_0 = 20$	0.036	0.024	0.022	0.020
$N_0 = 30$	0.073	0.022	0.019	0.017

Note: The number of subjects in the simulation is 40. N_0 indicates the number of subjects with no break in the simulation. The MCMC iteration for each simulation is 1,000 after 1,000 burnin iterations. Prior distributions are $\boldsymbol{\beta} \sim \mathcal{N}(0, 2), \sigma^2 \sim \mathcal{IG}(2, 1)$, and $p_{ii} \sim \mathcal{B}eta(T/20, 0.1)$.

Table 3: Shifting Relationships Between Inequality and Economic Development

	Mean	St.Dev	Credible Interval (95%)	
No Break Estimates				
α_0	-1.080	0.458	-1.974	-0.193
Lagged Top 1	-0.012	0.007	-0.025	0.001
Growth	1.803	0.253	1.315	2.296
Lagged log GDP	0.120	0.040	0.042	0.198
σ^2	0.375	0.023	0.333	0.423
D	0.690	0.275	0.340	1.395
Regime-specific Estimates				
	Regime 1: 1923 - 1952			
α_0	-0.407	1.148	-2.659	1.840
Lagged Top 1	-0.048	0.022	-0.091	-0.005
Growth	0.907	0.604	-0.282	2.097
Lagged log GDP	0.101	0.124	-0.140	0.342
σ^2	0.777	0.080	0.637	0.948
D	0.690	0.270	0.346	1.368
	Regime 2: 1953 - 1986			
α_0	0.657	0.222	0.217	1.087
Lagged Top 1	-0.018	0.005	-0.027	-0.009
Growth	-0.093	0.247	-0.579	0.388
Lagged log GDP	-0.064	0.024	-0.112	-0.016
σ^2	0.074	0.007	0.062	0.089
D	0.669	0.261	0.333	1.331
	Regime 3: 1987 - 1998			
α_0	1.975	2.160	-2.290	6.162
Lagged Top 1	0.049	0.017	0.014	0.081
Growth	0.505	1.680	-2.778	3.759
Lagged log GDP	-0.222	0.228	-0.668	0.230
σ^2	0.291	0.049	0.209	0.402
D	0.696	0.283	0.342	1.417

Note: The dependent variable is the top 1 percent income share. Prior distributions are $\boldsymbol{\beta} \sim \mathcal{N}(0, 10), \sigma^2 \sim \mathcal{IG}(1, 1), \mathbf{D} \sim \mathcal{W}(10, 0.1)$.

Table 4: The Effects of the GATT on Bilateral Trade Among Industrial Dyads

	Dyad-Year Fixed-Effects	Fixed-Effects HMM
Both participate in the GATT/WTO	0.749	0.106
Only one participates in the GATT/WTO	0.285	0.085
Reciprocal PTA	0.311	0.205
Nonreciprocal PTA	-0.047	0.064
GSP	0.016	0.033
Currency union	0.557	0.174
Log product real GDP	1.136	0.807
Dyad	594	
N	28971	

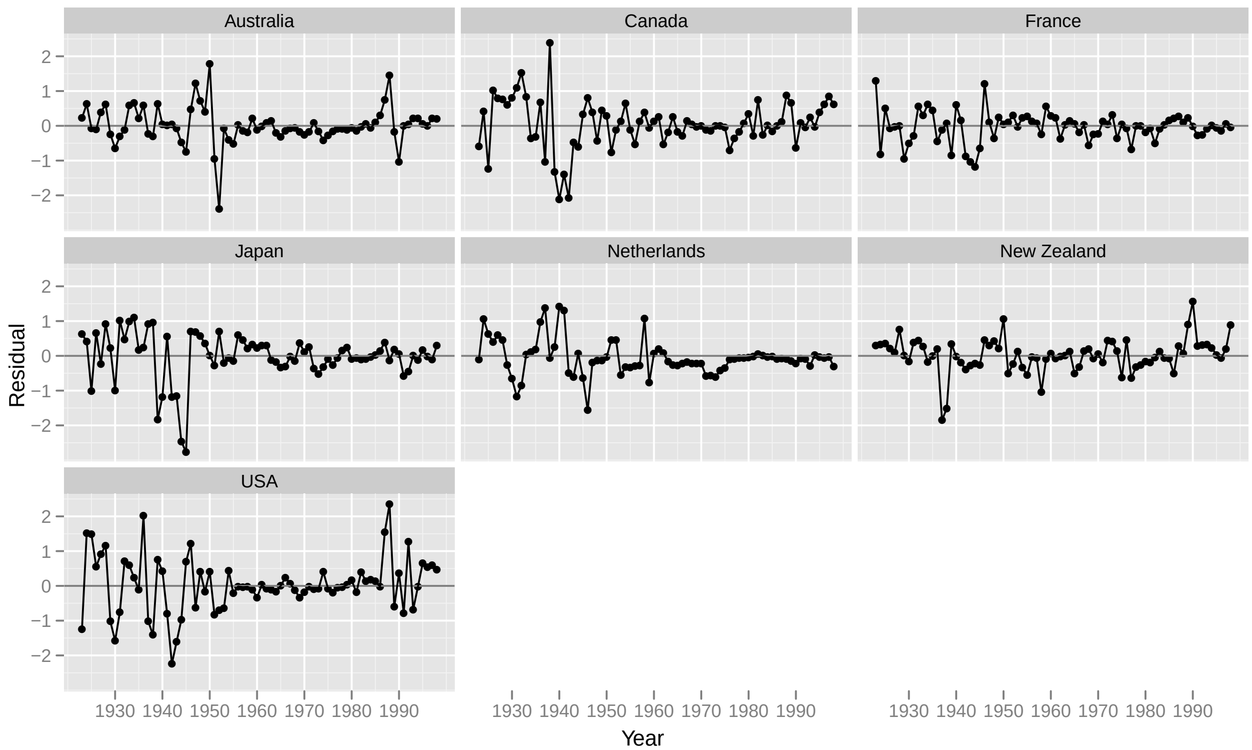
Note: The effects are computed by the formula $e^{\beta} - 1$. The unit of observation is the directed dyad-year. The dependent variable is the natural log of imports (in 1967 US dolloars). The colonial orbit dummy variable is dropped as no industrial dyad belongs to colonial orbit in the sample. The dyad and year fixed-effects estimates are from ordinary least squares (OLS) regression using lm function in R.

Table 5: The Effects of the GATT on Bilateral Trade Among All Dyads

	GRT	Dyad and Regime Fixed-effects
Both participate in the GATT/WTO		
Both formal members	0.298*	0.128*
	(0.012)	(0.010)
Both nonmember participants	0.428*	0.179*
	(0.027)	(0.023)
Formal member and	0.350*	0.126*
nonmember participant	(0.013)	(0.011)
Only one participates in the GATT/WTO		
Formal member	0.173*	0.035*
	(0.010)	(0.009)
Nonmember participant	0.155*	0.005
	(0.015)	(0.013)
Reciprocal PTA	0.326*	0.285*
	(0.007)	(0.007)
Nonreciprocal PTA	−0.062*	−0.045*
	(0.010)	(0.010)
GSP	−0.120*	−0.019*
	(0.006)	(0.005)
Currency Union	0.509*	0.363*
	(0.027)	(0.026)
Colonial Orbit	0.818*	0.623*
	(0.030)	(0.028)
Log product real GDP	0.633*	0.560*
	(0.003)	(0.002)
R^2	0.341	0.261
adj. R^2	0.340	0.261
Resid. sd	0.917	0.772
N	381,656	381,656

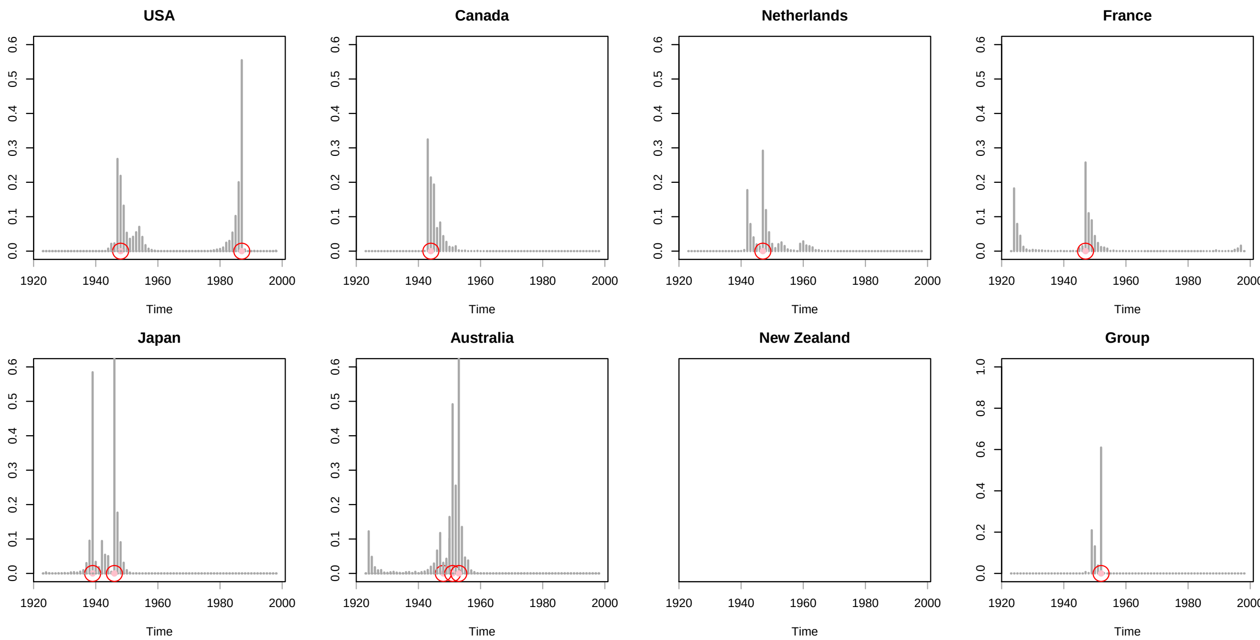
Note: Estimates from ordinary least squares OLS regression. For the fixed-effect estimation, the data sets are demeaned at the dyad and year level (GRT) and at the dyad and dyad-specific hidden regimes (Dyad and Regime Fixed-effects). The unit of observation is the directed dyad and the dependent variable is the natural log of imports measured in 1967 U.S. dollars. Standard errors in parentheses. * indicates significance at $p < 0.05$

Figure 1: Country-Specific Residuals from the Random Effects Error Correction Model of Inequality and Economic Development



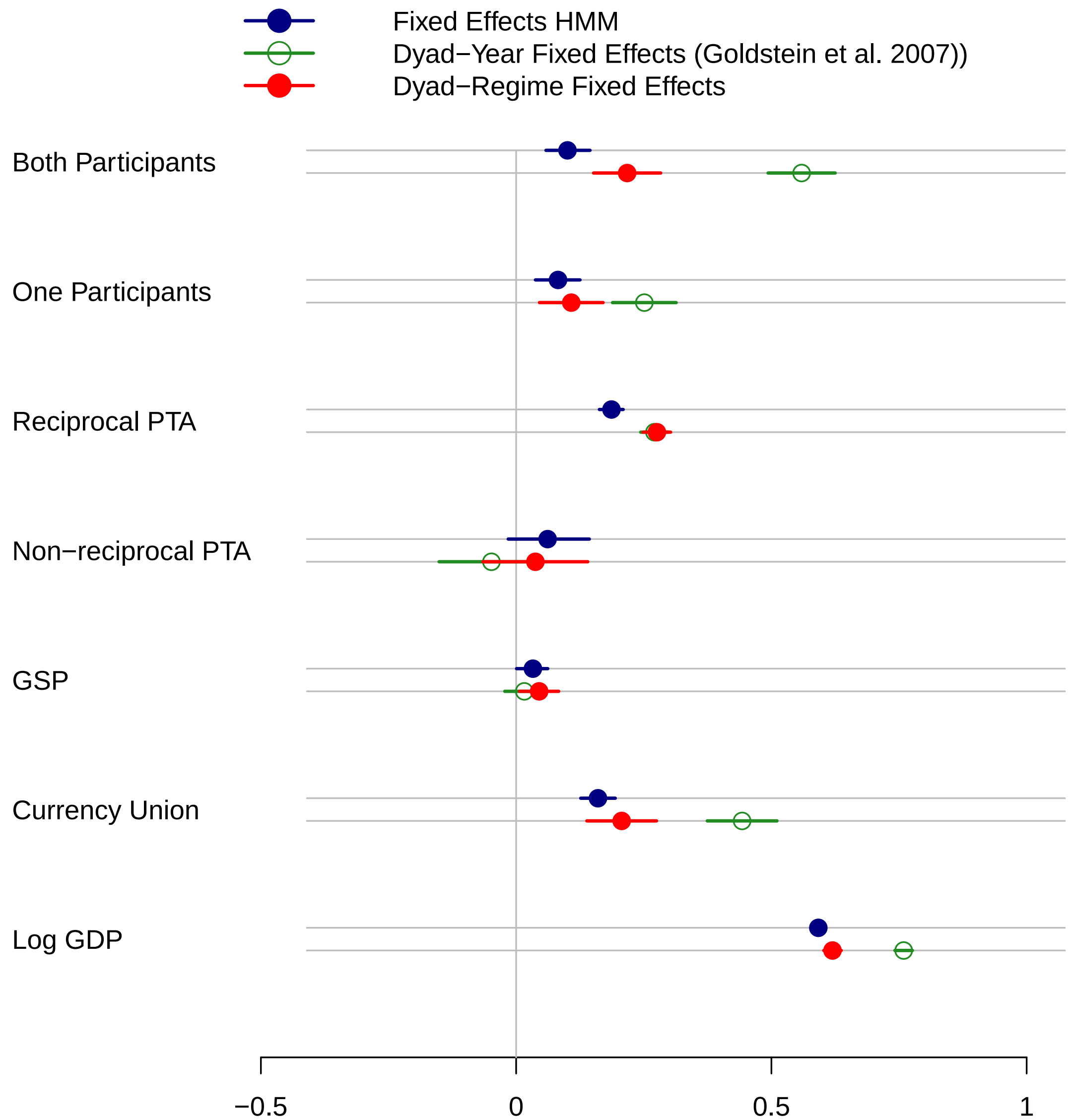
Note: The dependent variable is the top 1 percent income share. Included regressors are lagged top 1 percent share, economic growth rate, and lagged log GDP per capita. The top 1 percent data set is from Leigh (2007) and GDP per capita is from Maddison (2010). The model estimation was done using lme4 package in R.

Figure 2: Country-Specific Hidden Regimes in the Residuals of the Random Effects Error Correction Model of Inequality and Economic Development



Note: Large circles at the bottom indicate estimated break points and vertical bars are posterior probabilities of breaks. The data for the analysis are country-specific residuals from the random-effects error correction model using top 1 percent income share between 1923 and 1998. Included countries are Australia, Canada, New Zealand, USA, France, Japan, and The Netherlands. The top 1 percent data set is from Leigh (2007) and GDP per capita is from Maddison (2010). Employed prior distributions are $\beta \sim \mathcal{N}(0, 2)$, $\sigma^2 \sim \mathcal{IG}(4, 4)$, and $p_{ii} \sim \mathcal{B}eta(38, 1)$.

Figure 3: Comparison of Parameter Estimates from Different Fixed-Effects Models (Industrial Dyads Only)



Note: The unit of observation is the directed dyad-year. The dependent variable is the natural log of imports (in 1967 US dolloars). The dyad and year fixed-effects estimates are from ordinary least squares (OLS) regression using lm function in R.