The Evolution of Purchasing Power Parity

Collin Rabe
University of Richmond *

Andrea Waddle
University of Richmond †

A large body of literature in international finance has attempted to estimate the speed of convergence between countries' aggregate price indices to those levels predicted by purchasing power parity (PPP). This paper takes a novel approach by considering how this speed of convergence itself has evolved over time. Using a dynamic common correlated effects (DCCE) framework from Chudik and Pesaran (2015) applied to a panel of countries' real exchange rates over the years 1960–2015, we find an average half-life of around 3 years. More interestingly, we also show that the estimated half-life fell by about 1.5–3 years over the course of the past five decades, suggesting that the so-called PPP puzzle (Rogoff, 1996) may become an antiquated concern in the future. Our results also serve to contextualize past estimates by demonstrating the degree of sample selection sensitivity. Furthermore, we propose explanations for the observed increase in price re-calibration speed, focusing primarily on the increasingly tradable nature of the composition of the U.S. consumer price index (CPI). We build a measure of tradability of the CPI and show that, despite an increase in the proportion of services in the average consumer's basket, the CPI has become more tradable over time, thus offering a potential explanation for the observed increase in adjustment speed.

Keywords: purchasing power parity, exchange rates, price indices, international finance

*Contact: crabe@richmond.edu
†Contact: awaddle@richmond.edu

1 Introduction

Understanding the behavior of prices across borders has long been fundamental to economists' thinking about international trade and exchange rate systems. Numerous studies have sought to estimate the speed at which prices between countries converge over time by examining the dynamic behavior of real exchange rates (RERs). These studies have differed both in terms of methodologies and sample data, resulting in a wide range of estimates. This paper contributes to this literature by providing new estimates of adjustment speed utilizing an up-to-date econometric framework, but moreover focuses on an important related question: how has the speed of international price convergence itself changed over time? In other words, do global market forces more quickly push economies' prices toward purchasing power parity (PPP) today than in the past? Has the globalization of recent decades made the theory of PPP a better or worse paradigm for the workings of the modern world market?

We present empirical evidence that the speed of international price convergence has been *increasing* – or, equivalently, the time to convergence has been *decreasing* – over the past 55 years. As measured in half-lives, we find that the average rate of PPP convergence over the sample period of 1960–2015 is about 3 years, putting it in line with previous estimates from the literature. However, this rate seems to have fallen over time from a high point at the beginning of the sample to its current nadir, representing a decline of about 1.5–3 years over the entire sample period. These findings suggest that concerns about the PPP puzzle (see Rogoff, 1996) may ultimately become a thing of the past, if half-lives continue to fall into the future. Our results also serve to help contextualize the previous literature by visualizing the degree of sample selection sensitivity inherent in this type of estimation.

For an overview of the early empirical work on PPP, see Froot and Rogoff (1995). Most of the analyses of this period focused on improving the power of the estimates by increasing the sample size along one of two dimensions: the number of countries in the cross-section or the number of years for a particular pair of countries. Manzur (1990) finds among 7 industrialized countries an average half-life of 5 years, and Fung and Lo (1992) similarly considers 6 industrialized countries, finding a half-life of 6.5. Frankel and Rose (1996) famously examines a large cross-sectional panel of data from 150 countries and finds an average half-life of 4 years. Papers with long time dimensions include Abuaf and Jorion (1990) [8 currency pairs, 1901–1972, half-life of 3.3 years], Frankel (1986) [USD/GBP, 1869–1984, 4.6 years], Edison (1987) [USD/GBP, 1890-1978, 7.3 years], Johnson (1990) [USD/CAD, 1914–1986, 3.1 years], and Lothian and Taylor (1996) [USD/GBP, 1791–1990, 6 years; FRF/GBP, 1803–1990, 3 years]. Rogoff (1996) summarizes all of this work as having achieved a "remarkable consensus" on an observed RER half-life of 3–5 years.

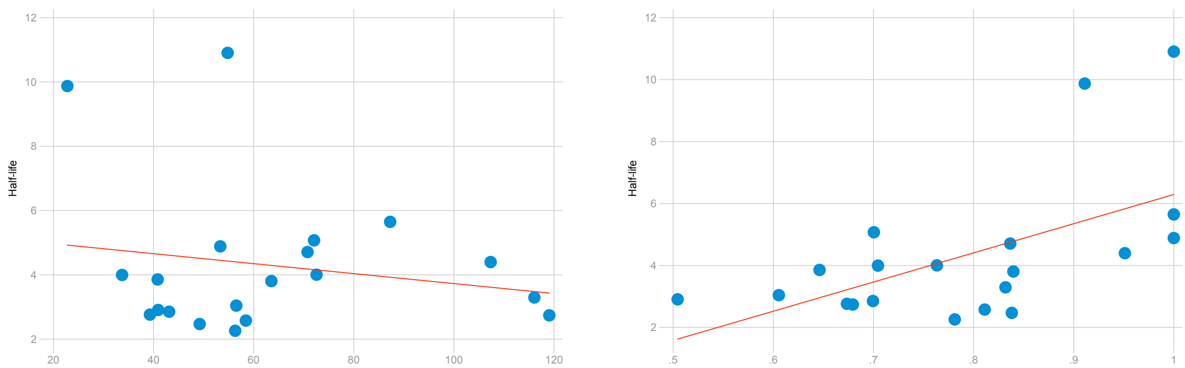
There are numerous reasons for the variation in these estimates. Part of the difficulty stems from the fact that the half-lives are derived from nonlinear transformations of the parameters actually being estimated, making them extremely sensitive to even minuscule differences in the models' point estimates. Additionally, small differences in the samples used (even for the same groups of countries) can generate very different results. In our analysis, we show how a change in sample of only two years' worth of observations can result in the estimated half-life changing by as much as one full year. Furthermore, many of these early studies suffered from various forms of statistical bias. Dynamic panel bias in fixed effects models (*a la* Nickell, 1981) can be substantial for studies covering a finite number of time periods, even when the number of countries involved is very large.

More recent work has tried to update the 3–5 year consensus by employing more advanced econometric techniques. Cheung and Lai (2000) use an impulse response analysis to demonstrate a half-life of 2–5 years for industrialized countries and under 3 years for developing countries. Taylor (2001) focuses on the temporal aggregation of observed data, as well as the possibility that there are nonlinearities in the real exchange rate process, arguing that both issues may causes upward bias in the estimation of half-lives. Imbs et al. (2005) demonstrate how failing to account for dynamic heterogeneity in the determination of real exchange rates may also introduce significant upward bias. Choi et al. (2006) addresses the issues of Nickell bias, time aggregation, and heterogeneous dynamics by deriving an analytical bias correction formula, and they estimate an average half-life of 3.5 years among 20 OECD countries.

Many of the studies dealing with the issue of heterogeneous dynamics implement approaches based on the pioneering work of Pesaran and Smith (1995), which provided a theoretical foundation for the use of mean group estimators. Subsequent work, particularly Pesaran (2006), highlighted the importance of simultaneously controlling for the impact of cross-sectional dependence using a common correlated effects (CCE) framework. This was later extended to a dynamic common correlated effects (DCCE) framework in Chudik and Pesaran (2015) to allow for the consistent estimation of a CCE model that includes a lagged dependent variable, such as is commonly used in the estimation of RER half-lives. This DCCE estimator serves as the foundation for our methodological approach.

The primary motivation for our work is that all of these studies operate under the assumption that the underlying relationships are constant over the whole time horizon. Especially when considering 100 years or more of international economic history, this assumption of structural consistency seems unlikely. Using Monte Carlo simulations, we demonstrate in Section 3.2.3 that when the presumed-constant parameter being estimated actually follows a dynamic trend, standard econometric models can suffer from significant bias in the estimation of that parameter's mean value over the time horizon. By acknowledging the potential for change in the behavior of RERs over time and attempting to estimate it, we hope to provide a fuller empirical picture of the relevance of PPP throughout history and some basis for forecasting its role in the future.

Perhaps the closest previous work to this paper comes from Hegwood and Papell (1998), Astorga (2012), and Balli et al. (2014), who investigate the behavior of real exchange rates vis-a-vis PPP-implied levels in the presence of structural breaks. All studies find that incorporating apparent structural breaks into their econometric models has notable effects on the estimated mean-reversion behavior of exchange rates. Our paper differs from and extends this previous work in a number of important ways: 1) we consider a broader panel of countries, 2) we control for common correlated effects using a theoretically consistent estimator, and 3) we approach the modeling of structural dynamics in a fundamentally different way. While these papers allowed for structural change in the behavior of RERs, the focus was on identifying and accounting for *discrete* breaks. While there certainly are major economic events that could introduce sudden structural shocks in the context of exchange rate determination (e.g. changes in exchange rate regimes), our framework postulates that structural change primarily is driven *continuously* and gradually over decades by the changing dynamics in the forces of trade, technology, communication, markets, capital flows, or "globalization" in general.



(a) Trade Openness

(b) Capital Account Openness

Figure 1: Half-lives (in years) of real exchange rate convergence in OECD sample countries (see Section 2), 1960-2015, using individual coefficient results from estimation of DCCE model (see Section 3). Trade openness is the sum of exports and imports (as a percentage of GDP), and capital account openness comes from the normalized Chinn-Ito index (most open = 1), both measured as averages over the whole sample period.

When imagining a world in which the half-life of RER shocks has been shifting over time, it's unclear what our ex ante expectation should be, especially considering the impacts of globalization on international price behavior. The concept of PPP is almost deceptive in its intuitiveness, being based as it is upon the simple idea of the law of one price (LOOP), i.e. market participants will move quickly to eliminate arbitrage opportunities that arise when identical goods are priced differently in different locations. Of course, local prices could be influenced by a multitude of additional factors, including product differentiation, branding, consumption preferences, market composition, etc. However, if we suspect many goods to be fungible, homogeneous, and/or incurring low transportation costs, then we might reasonably expect prices to realign more quickly among those countries that trade more extensively, and therefore bring those international arbitrage forces more strongly to bear on their economies. Moreover, actors in countries that trade more extensively may be reasonably expected to have faster access to information identifying price discrepancies that can be exploited. Figure 1a presents a simple regression of countries' price convergence half-lives on their openness to international trade. The observed relationship is expectedly negative, albeit somewhat weak. Thus, as we have observed a trend of greater and greater cross-border trade among countries in the post-war period, we might reasonably expect an increase in adjustment speeds to PPP.

Concomitantly, however, we must keep in mind that international prices have two components: the local currency price and the nominal exchange rate. Moreover, the nominal exchange rate does not move solely to eliminate trade arbitrage opportunities arising from international price discrepancies. In fact, the nominal exchange rate could never adjust unilaterally so as to perfectly achieve absolute PPP, as each economy's aggregate price index comprises hundreds of thousands of prices simultaneously requiring movements in conflicting directions so as to individually satisfy the law of one price. Instead, the nominal exchange rate is also buffeted by international financial forces with entirely different motivations. For example, hot money flows chasing higher interest rates and/or a reallocation of portfolio risk may push a nominal exchange rate in the opposite direction of that prescribed by PPP. Figure 1b presents a simple regression of price convergence half-lives on countries' degrees of openness to financial flows, as measured by the de jure index of capital account policies by Chinn and Ito (2006). We find that countries that are more open to financial flows tend to have slower convergence to PPP. This coincides with the finding in Bergin et al. (2017) that eurozone countries experienced an increase in price convergence speed after adopting the euro and thereby removing any interference stemming from nominal exchange rate fluctuations. Therefore, depending on how we interpret "globalization" and the relative sizes of its component influences, the sum effect on the speed of international price convergence could theoretically be positive, negative, or negligent.

By applying a DCCE estimator to 1) a single, "consolidated" regression model, and 2) a "rolling windows" framework, we find that the half-life of RER deviations from PPP have fallen substantially over the past several decades. We examine the theoretical predictions for the potential sources of these observed empirical results, exploring in particular the possibility that changes in the measurement itself of real exchange rates (via the tradable composition of US CPI) may have greatly contributed to them. It is important to note that the weights placed on various goods and services in the CPI change over time, primarily due to a change in the fraction of the consumer's income spent on that good or service. Therefore, the extent to which items in the CPI are traded is ambiguous, as trade lowers the price of traded goods, thus decreasing their relative weight in the CPI, even as the measured "tradabilty" of those items increases. In order to explore the tradability of the price index, we construct several measures of the "tradable CPI" which take into account both the changing weights of the items that comprise the CPI, as well as the extent to which they are traded. We find that the tradability of the CPI increased between 250 and 500% from 1970 to 2015. This dramatic increase in the tradability of the average basket of goods and services suggests that trade is responsible, at least in part, for the observed decrease in adjustment speed of deviations from PPP.

The paper proceeds as follows: We discuss our data set and its statistical properties in Section 2. Section 3 introduces our estimation strategy and provides some Monte Carlo simulations for motivation. Section 3.3 presents our empirical results in detail. Section 4 offers some explanations for the observed results, and explores in depth how the composition of CPI has changed over time. We conclude with some final observations and avenues for further research in Section 6.

2 Data

To conduct our analysis, we create a data set of annual observations of real exchange rates for a panel of countries vis-a-vis the US dollar according to the following definition:

$$q_{i,t} = s_{i,t} + p_{US,t} - p_{i,t} \tag{1}$$

where i and t are country and year subscripts, respectively, q represents the real exchange rate (an increase denoting a real depreciation from country i's perspective), s is the nominal exchange rate in local currency units per US dollar, and p represents national price levels, all expressed in logarithmic terms.[1] As noted in previous work, the choice of base country can have minor impacts on empirical results in certain contexts. However, throughout this paper we focus exclusively on RERs defined using the US as the base country for both the sake of brevity in reporting results and to more directly tie into our detailed discussion of the composition of US CPI in Section 4.[2]

We use nominal exchange rate data from the International Monetary Fund's International Financial Statistics (IFS) and make use of CPI data pulled from the IFS, Organization for Economic Development (OECD), and the World Bank's World Development Indicators (WDI) databases as proxies for aggregate price levels.[3] Data starts in 1960 and ends in 2015, although country representation is not consistent across time, especially in early years. As such, we drop any countries that lack continuous observations over this horizon, resulting in RER series for 66 developed and developing countries across the world.

In our baseline estimations, we limit the panel to a set of 20 industrialized OECD member countries,[4] as used in Choi et al (2006), which provides a useful point of reference. In fact, the majority of papers in this line of literature have focused on similar groups of developed

[1]Some recent studies, such as Taylor (2002) and Bergin et al. (2017), have utilized time de-trended RERs in the context of PPP analyses. This has been justified as capturing the Balassa-Samuelson effect. We use the RER in levels out of concern that de-trending may remove important information as it pertains to the dynamics of RER behavior over time, and moreover because the high-income countries in our baseline "OECD" group have not experienced a significant amount of productivity catch-up relative to each other.

[2]Results using alternative countries as the base are available from the authors upon request.

[3]Due to differences in data sources and collection methods, the coverage of CPI data provided by each these databases varies significantly. For our purposes, the most important thing is that the data series are internally consistent across time for each country, so we choose the longest continuous series available out of the three databases on a country-by-country basis. Alternative selection methods don't greatly affect our results, since in cases where there is overlap across data sets the vast majority of differences are negligible ($< 1\%$).

[4]This "OECD" group comprises Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom.

economies, so we wanted to allow for easy comparison across studies. Moreover, as a group of the most developed and open economies in the world, both in terms of trade and financial flows, these countries should most clearly exhibit the influences of the law of one price. We then extend our analysis to the full panel of 66 countries with available data, which display a much greater degree of economic diversity.[5]

We next consider some important statistical characteristics of our data set. First, following the procedure in Pesaran (2015), we test for evidence of cross-sectional dependence across countries, the presence of which could introduce substantial bias into our regression results if not taken into account. We obtain test statistics of $CD = 54.1$ and $CD = 217.7$ for our "OECD" and all-countries panels, respectively, which correspond to p-values of ≤ 0.000, meaning we can reject the null hypothesis of "weak cross-sectional dependence" in favor of "strong" dependence. Because of this empirical evidence, our preferred estimation framework will incorporate additional controls to compensate for the evident dependence.

Next, we consider the issue of stationarity. Since the focus of this paper is on how RER convergence speeds have changed over time, we will implicitly be assuming that the underlying series are in fact stationary. There already exists a vast literature examining this question in particular. See, for example, work by Papell (2006) which finds increasing evidence of RER stationarity for a similar panel of OECD countries during 1973–1998 using a panel ADF test, as well as Holmes et al. (2012). To ensure there is nothing abnormal about our assembled data, we also run a CADF test on our full time sample, following Pesaran (2007)'s simple panel unit root test in the presence of cross-sectional dependence, which calculates a test statistic based on the mean of t-stats from Dickey-Fuller tests on each observational unit in the panel. The null hypothesis is that all series are non-stationary. Allowing for a constant term and one lag, we reject the null hypothesis with $t = -3.816$ and $t = -2.242$ for the "OECD" and all-countries panels, respectively, again both corresponding to p-values of ≤ 0.000. In the following section, we lay out our estimation strategy in detail.

3 Estimation

3.1 Static Models

We start by considering a simple AR(1) model for the real exchange rate:

$$q_{i,t} = \alpha_i + \rho \cdot q_{i,t-1} + e_{i,t} \tag{2}$$

[5]The full set of countries includes the set of 20 "OECD" countries listed above, as well as Argentina, Bolivia, Burkina Faso, Chile, Colombia, Costa Rica, Cote d'Ivoire, Cyprus, the Dominican Republic, Egypt, El Salvador, the Gambia, Guatemala, Haiti, Honduras, Iceland, India, Iran, Israel, Jamaica, Kenya, Luxembourg, Malaysia, Malta, Mexico, Morocco, Myanmar, Nigeria, Pakistan, Panama, Paraguay, Peru, the Philippines, Samoa, Singapore, South Africa, South Korea, Sri Lanka, Sudan, Suriname, Thailand, Trinidad and Tobago, Turkey, Uruguay, and Venezuela.

for $i = 1 \ldots n$ and $t = 1 \ldots T$, with a common autoregressive coefficient ρ across all observational units and an idiosyncratic error term e. In this context, if q is a stationary process, then its long-run mean can be expressed as $\mu_i = \alpha_i/(1-\rho)$. Note that if the law of one price held perfectly between countries for every good, then absolute PPP would dictate that $\mu_i = 1 \; \forall \; i$. However, because of differences in the baskets of goods and services that define CPI indices in every country, we must allow for heterogeneity in the PPP-implied long-run means. The half-life of reversion to these mean values in response to a random shock can be expressed as:

$$\lambda = \frac{\log 0.5}{\log \rho} \tag{3}$$

Thus, measurement of the speed of international price convergence to PPP levels simply requires an accurate estimate of ρ.

Unfortunately, empirical estimation of ρ is problematic for a number of reasons. First, when $n = 1$, OLS estimation is biased because of the violation of strict exogeneity of the errors across time. It is, however, consistent as $T \to \infty$, although T is unfortunately often fairly small in applied macroeconomics. One way to improve the accuracy of the estimation is by expanding the number of observational units (i.e. $n > 1$), but for any finite T this introduces Nickell bias that can be quite substantial, even when $n \to \infty$. This is because using a least-squares dummy variables (LSDV) estimator (equivalent to OLS using time-demeaned observations) in an autoregressive panel context means that the time-demeaned values of the lagged dependent variable on the right-hand side of the regression equation depend on future-period values, which are in turn defined by and hence correlated with current-period shocks. These issues are, however, well-known and numerous methodologies have been developed to address them, either through analytical derivations of bias-correcting formulae (Choi et al., 2006), bootstrap resampling (Phillips and Sul, 2007), or half-panel jackknife techniques (Dhaene and Jochmans, 2015; Chudik et al., 2018).

Third, more recent research utilizing large panels has emphasized the bias-inducing effects of error cross-sectional dependence. A prominent approach to dealing with this dependence assumes a multifactor error structure, such that the error term in (2) is further defined as

$$e_{i,t} = \boldsymbol{\gamma}_i \boldsymbol{f}_t + \epsilon_{i,t} \tag{4}$$

where $\boldsymbol{f}_t$ represents a finite number of unobserved (and possibly serially correlated) common factors that influence each observational unit idiosyncratically depending on the value of $\boldsymbol{\gamma}_i$, and ϵ is a true white-noise error term. Standard LSDV estimation of a model with this type of error structure is no longer consistent. Pesaran (2006) introduced a common correlated effects (CCE) estimator that retains consistency in a model with unobserved factors and exogenous regressors. Chudik and Pesaran (2015) then extended this to a dynamic common correlated effects (DCCE) estimator that further achieves consistency when lagged observations of the dependent variable are included, by adding lags of the cross-sectional averages of the regressors to the regression equation.

Fourth, Pesaran and Smith (1995) point out that pooled estimators incorrectly assuming homogeneous coefficients across countries can suffer from substantial bias in the estimation of the average effect, even when N and T are large, due to induced serial correlation in the disturbances. Supposing that the autoregressive coefficient in (2) is not actually common across all countries, we then generalize the model by writing:

$$\begin{aligned} q_{i,t} &= \alpha_i + \rho_i \cdot q_{i,t-1} + e_{i,t} \\ \rho_i &= \rho + \nu_i \end{aligned} \tag{5}$$

where ν_i are independent and identically distributed random deviations. The estimation approach provided by Chudik and Pesaran (2015) simultaneously addresses the confounding effect of heterogeneous coefficients in addition to the cross-sectional dependence described above. Thus, throughout the paper, our "DCCE" results are reporting mean group estimates that account for potential heterogeneity as in (5).

While all of the issues discussed above affect the accurate estimation of the half-life of shocks to RERs, the primary focus of this paper is on the possibility that the half-life itself exhibits changes over time. In the next section, we consider how to modify the framework above to capture coefficient dynamics and how this further complicates the estimation of PPP half-lives.

3.2 Dynamic Models

Consider the estimation of the AR(1) process when the coefficient has its own dynamics:[6]

$$q_{i,t} = \alpha_{i,t} + \rho_t \cdot q_{i,t-1} + e_{i,t} \tag{6}$$

and therefore by extension:[7]

$$\lambda_t = \frac{\log 0.5}{\log \rho_t} \tag{7}$$

By virtue of the new time subscript, a decreasing value of ρ_t over time would generate a progressively lower value for the half-life λ_t, since we continue to assume the stationarity of $q_{i,t}$ and therefore $\rho_t < 1$. Furthermore, note that stationarity and the very concept of a half-life necessitate the existence of some constant long-run mean value to which the process reverts. Thus, while the half-life only depends on the value of ρ_t, we must also add a time subscript to the drift term, $\alpha_{i,t}$, that "compensates" for the changes in ρ_t and ensures the mean value of μ_i is constant in the long-run.

Accurate estimation of this model is difficult because standard econometric techniques assume

[6] To simplify and focus the exposition on the impact of dynamic changes in the autoregressive coefficient, we maintain the assumption of coefficient homogeneity across countries in this section, while allowing for heterogeneity in the estimation results presented in Section 3.3.

[7] Technically, this expression for half-life is only correct when assuming ρ_t will remain constant in all future periods following the initial shock from which the process is reverting. See Section 3.3 for more discussion.

the constancy of the underlying parameters of the data generating process. To illustrate the implications of estimating a dynamic-coefficient process using a static-coefficient model, we present the following Monte Carlo analysis. We simulate a panel of AR(1) processes according to (6) using decreasing values of ρ_t following a linearly deterministic path and apply a standard LSDV estimator to the generated time series. *A priori*, one might hope that the LSDV estimator would at least provide an accurate estimate of the mean value of the coefficient over the whole time horizon, i.e. $\bar{\rho} \equiv \frac{1}{T}\sum_{t=1}^{T}\rho_t$. However, as seen in Table 1, there is substantial bias, whose sign and magnitude depends on the size of the time dimension. Furthermore, increasing the cross-sectional dimension does not have a profound effect on the size of this bias. For context, a bias of -0.01 corresponds to a difference in half-life of about 8 months when $\bar{\rho} = .9$ and 0.5 months when $\bar{\rho} = .6$.

Monte Carlo: LSDV Estimation of $\bar{\rho}$

	Bias (×100)					RMSE (×100)				
$(n,\ T)$	25	50	75	100	150	25	50	75	100	150
10	-6.59	-2.79	-0.87	0.95	6.03	8.83	4.95	3.49	3.19	6.71
20	-6.49	-2.70	-0.80	1.01	6.08	7.72	3.94	2.51	2.36	6.42
50	-6.32	-2.65	-0.75	1.06	6.11	6.85	3.22	1.69	1.74	6.26
100	-6.33	-2.61	-0.76	1.06	6.13	6.60	2.92	1.32	1.44	6.20

Table 1: Based on LSDV estimations of 5,000 simulations of a panel of AR(1) processes, generated with $\bar{\rho} = 0.6$ (mean of ρ_t over T), $d\rho = -0.005$ (period-to-period change in ρ_t), $\mu_i \sim N(1, 0.01)$ (distribution of means), and $e_{i,t} \sim N(0, 0.0225)$ (distribution of stochastic shocks).

To be clear, the focus of this paper is not so much on estimating the *level* of the autoregressive coefficient (and the corresponding half-life, by extension), but rather the *change* in the coefficient over time. To this end, we next consider two approaches to estimating the change in the coefficient in the dynamic AR(1) model in (6): 1) a "rolling windows" model, and 2) a "time-trending regression" model.

3.2.1 Rolling Windows Model

The first approach we consider does not directly model the dynamic behavior of ρ_t, but rather applies the static framework in (2) to different time subsets, or "windows," in an effort to reveal the underlying dynamics. The use of rolling overlapping windows in applied empirical analyses has been employed in many fields and for a variety of purposes, e.g. forecasting, identifying potential structural breaks, and revealing the possible time-inconsistency of the underlying parameters. Examples from the international finance literature include Cheung et al. (2005), Vigfusson et al. (2009), and Manzur and Chan (2010).

For our purposes, the rolling windows approach involves repeatedly estimating the following model (with or without a multifactor error structure):

$$q_{i,t} = \alpha_i + \rho_j \cdot q_{i,t-1} + e_{i,t}, \quad t = j, \ldots, j + w - 1 \tag{8}$$

for $j = 1, \ldots, T - w + 1$, where w is the number of periods in each subset of the total number of sample periods, T. Thus, we obtain a pair of parameter estimates of $\hat{\alpha}_{ij}$ and $\hat{\rho}_j$ for each window j. Then, in a second stage, we collect the $T - w + 1$ estimates of ρ_j and use them to estimate the following:

$$\hat{\rho}_j = \rho_0 + d\rho \cdot j + \epsilon_j \tag{9}$$

where $d\rho$ quantifies the change in ρ over time. Technically, this second-stage regression is estimating the change in ρ from *window-to-window*, not *year-to-year*, but since we're using overlapping windows with no gaps, these two measures are equivalent under the assumption of a linear adjustment model. Furthermore, note that even though this model is effectively re-estimating a new "long-run" mean for each country in every window, this isn't necessarily a rejection of a broader notion of long-run stationarity across all windows. That is, the estimated values of α and ρ are still free to imply a consistent mean across windows. Moreover, the reason we allow for country-specific means in the first place is because of the idiosyncratic definitions of CPI measures across countries, and these definitions themselves are not necessarily static across time (see Section 5), thus necessitating updated mean values to which the series should be observed reverting.

Again, our primary objective is obtaining an accurate estimate of the change in half-lives across the sample period, rather than accurate estimates of the levels of the half-lives themselves. In other words, even if our estimates of the coefficient *levels* are biased, as long as they are all biased in the same way then the estimated *change* will be unbiased. For example, if the true values are $\rho_t = .8$ and $\rho_{t-1} = .7$, but our estimation procedure introduces positive bias such that we estimate $\hat{\rho}_t = .9$ and $\hat{\rho}_{t-1} = .8$, we still conclude that the change in the coefficient $d\rho = 0.1$ in both cases. Therefore, a potential advantage of this estimation approach is the ability to cancel out whatever biases might exist through differencing, conditional on the biases being somewhat consistent in terms of both sign and magnitude over time. Of course, we can't know for sure in practice that the bias in our estimates is consistent across windows, but it does seem somewhat reasonable when applying the same estimation procedure to adjacent windows that only differ by two observations.

When fitting the regression equation in (8) for each window, we employ two different estimators: 1) a traditional fixed effects or least-squares dummy variables "LSDV" estimator, and 2) a "DCCE" estimator to acknowledge the possibility that the error term in (8) is best represented by the unobserved multifactor error structure in (4). Once the coefficient estimates for each window are obtained, the second-stage regression is performed the same way in both cases. Empirical results using both estimators are presented in Section 3.3.

3.2.2 Time-Trending Regression Model

We next consider the employment of a "time-trending regression" model. This approach essentially starts by assuming a functional form for the dynamics of the parameters over time, then substitutes these into the basic AR(1) framework in (6) to create a new consolidated model that can be applied to the entire data set. See Brown et al. (1975) for an early comparison of this methodology to a rolling windows approach. In our context, suppose that the linear relationship outlined in the second-stage regression in (9) is basically correct. More precisely, suppose we start by assuming that the autoregressive coefficient follows a deterministic path defined by:[8]

$$\rho_t = \rho_0 + d\rho \cdot t \tag{10}$$

where ρ_0 represents the initial value of the autoregressive coefficient, and $d\rho$ represents the constant change of the coefficient from one period to the next, which is ultimately our parameter of interest. If we apply the model to the data and find $d\rho$ to be zero, then we can conclude that the underlying process driving RER behavior is most likely time-invariant.

The next step is to substitute this deterministic expression into the RER dynamics equation in (6), which yields:

$$\begin{aligned} q_{it} &= \alpha_{i,t} + \rho_t \cdot q_{i,t-1} + e_{i,t} \\ &= \alpha_{i,t} + (\rho_0 + d\rho t)q_{i,t-1} + e_{i,t} \\ &= \alpha_{i,t} + \rho_0 \cdot q_{i,t-1} + d\rho \cdot t \cdot q_{i,t-1} + e_{i,t} \end{aligned}$$

However, at this point, the model violates the assumption that q_{it} is stationary. To see this, note that weak stationarity implies that we can write:

$$E[q_{i,t}] = E[q_{i,t-1}] = \mu_i$$

So, if we take the expectation of the consolidated equation above, we obtain:

$$\begin{aligned} E[q_{i,t}] &= E[\alpha_{i,t} + \rho_0 \cdot q_{i,t} + d\rho \cdot t \cdot q_{i,t-1} + e_{i,t}] \\ \mu_i &= \alpha_{i,t} + \rho_0 \mu_i + d\rho \cdot t \cdot \mu_i \\ \mu_i &= \frac{\alpha_{i,t}}{1 - \rho_0 - d\rho \cdot t} \end{aligned}$$

which is potentially a contradiction, because the stationary mean μ_i should be time-invariant, but the right-hand side of this equation clearly depends on time, both directly in the denominator and indirectly through the behvior of $\alpha_{i,t}$. To correct this, we must also assume that the value of $\alpha_{i,t}$ adjusts in tandem with the deterministic path of ρ_t in order to maintain a constant stationary mean over time. That is, we rearrange the expression above to obtain the path for

[8]Importantly, the assumed functional form here does not include an error term, as in (9). Including a stochastic term would be equivalent to assuming the coefficient is defined by a unit-root process, from which we could not hope to meaningfully identify any parameters.

$\alpha_{i,t}$:

$$\alpha_{i,t} = (1 - \rho_0 - d\rho \cdot t)\mu_i$$

and then substitute this into the full regression equation above to obtain the following fully-consolidated "time-trending regression" model:

$$\begin{aligned} q_{i,t} &= \alpha_{i,t} + \rho_0 \cdot q_{i,t-1} + d\rho \cdot t \cdot q_{i,t-1} + e_{i,t} \\ &= (1 - \rho_0 - d\rho \cdot t)\mu_i + \rho_0 q_{i,t-1} + d\rho \cdot t \cdot q_{i,t-1} + e_{i,t} \\ &= \lambda_{0i} + \lambda_{1i} \cdot t + \rho_0 \cdot q_{i,t-1} + d\rho \cdot t \cdot q_{i,t-1} + e_{i,t} \end{aligned} \tag{11}$$

where $\lambda_{0i} \equiv (1 - \rho_0)\mu_i$ and $\lambda_{1i} \equiv -d\rho \cdot \mu_i$. This model now implies the following relationship:

$$\frac{\partial q_{i,t}}{\partial q_{i,t-1}} = \rho_0 + d\rho \cdot t$$

which clearly allows for the degree of persistence in RERs to adjust dynamically across time periods. In contrast to the rolling windows model, this approach only requires estimating the single equation in (11) using the entire sample of $t = 1, ..., T$. Again, we employ both LSDV and DCCE estimators in fitting the model, and present results using both in Section 3.3. In both cases, note that the i subscript on λ_{1i} requires the estimation of country-specific parameters for both the constant *and* the coefficient on the time trend.

3.2.3 Monte Carlo Comparison

In this section, we employ a Monte Carlo analysis to compare the estimation performance of the two approaches outlined in the previous sections. We start by building panels of n = 10, 20, 50, and 100 observational units, each with time horizons of T = 50, 75, 100, and 150 periods. For each panel, we generate $5,000$ simulations of processes defined by an AR(1) model with dynamic coefficients, as in (6). Each observational unit has an idiosyncratic mean, μ_i, that is randomly drawn from a normal distribution with a mean of 1 and standard deviation of 0.1. Stochastic shocks in each period, $e_{i,t}$, are drawn from a zero-mean normal distribution with a standard deviation of 0.15. The autoregressive coefficient is defined to follow a deterministic path such that its value decreases by $d\rho = -.005$ each period, with a mean value of $\bar{\rho} = 0.6$ across all T periods. Correspondingly, the drift term for each observational unit, $\alpha_{i,t}$, follows the deterministic path that will maintain its long-run mean value of μ_i. Using these simulations, we then apply both the time-trending regression model and the rolling windows model (using windows of 30 periods) to evaluate their performance in estimating $d\rho$.[9] Table 2 presents the simulated results in terms of biases and root-mean-square errors (RMSE).

With the exception of panels with small time dimensions ($T < 30$, not reported), the biases are all positive, implying that our estimators will tend to *understate* the degree of dynamic

[9]For brevity, the application of DCCE estimators is not included in this analysis.

Monte Carlo: Estimation of $d\rho$

(n, T)	Bias ($\times 10,000$) 50	75	100	150	RMSE ($\times 100$) 50	75	100	150
				Time-Trending Regression Model				
10	1.34	2.73	2.55	1.96	0.27	0.15	0.09	0.05
20	2.27	2.83	2.49	1.89	0.19	0.11	0.07	0.04
50	2.04	2.76	2.42	1.83	0.12	0.07	0.05	0.03
100	2.20	2.71	2.33	1.77	0.09	0.05	0.04	0.02
				Rolling Windows Model				
10	2.53	2.06	2.29	2.01	0.34	0.19	0.12	0.06
20	3.35	2.36	2.01	1.82	0.24	0.14	0.09	0.04
50	2.70	2.03	1.95	1.64	0.16	0.09	0.06	0.03
100	2.94	2.03	1.86	1.61	0.11	0.06	0.04	0.02

Table 2: Based on LSDV estimations of 5,000 simulations of a panel of AR(1) processes, generated with $\bar{\rho} = 0.6$ (mean of ρ_t over T), $d\rho = -0.005$ (period-to-period change in ρ_t), $\mu_i \sim N(1, 0.01)$ (distribution of means), and $e_{i,t} \sim N(0, 0.0225)$ (distribution of stochastic shocks). Rolling windows estimation performed with 30-year windows.

change in the estimated half-life when the deterministic trend is negative. This is an interesting point to keep in mind when interpreting our empirical results in Section 3.3. Note that the bias values have all been scaled up by a factor of 10,000, as the estimators are all reassuringly accurate across the board. Of course, the magnitudes will depend on the parameterizations of the simulations, particularly the variance of the shocks, but suffice it to say that under the chosen values one would have a rather difficult time clearly observing the change in ρ_t just by observing the generated processes with the naked eye. As a minor note, we also considered the application of "half-panel jackknife" adjustments to correct for estimation bias, *a la* Dhaene and Jochmans (2015) and Chudik and Pesaran (2015). However, in the context of our model with time-variant coefficients, we found in our simulations that the jackknife adjustment would simply push down the estimated values of $d\rho$, thereby either reducing or exacerbating the bias depending on how accurate the estimate was before the adjustment. Since the unadjusted estimates were fairly accurate to begin with, and the sign of the bias is unambiguous, we deemed the application of jackknife methods too problematic and do not include them in our analyses.

With regards to cross-sectional size, increasing the number of observational units does improve the accuracy of the estimation, but only when the time dimension is sufficiently large, e.g. $n > 20$ and $T > 100$. When n is small, more observational units introduce more variation without necessarily providing more opportunity to accurately identify the underlying behavior of the series. Furthermore, in the case of the time-trending regression model, having a larger time dimension does not always reduce bias, but it does always improve the RMSE because of reductions in the variance of the sampling distribution. In contrast, the bias of the rolling windows model is generally decreasing as the time dimension increases, and in fact we find that it is possible to effectively reduce the bias to zero with a sufficiently high T and a correct corresponding window size. While this makes the rolling windows approach seem preferable in some respects, it unfortunately always suffers from relatively larger variances in its sampling distribution and therefore higher RMSEs. In instances where the sample size is sufficiently large along both dimensions, however, there is very little difference in the performance of the two estimation models. Therefore, we present results from the use of both approaches when applying them to the actual data in the next section.

Even when the sample size is not sufficiently large to eliminate all of the difference in accuracy between the two models, the rolling windows approach still has one important advantage: it allows for an easy way to visualize the dynamics. In other words, a drawback of the time-trending regression model is that the accuracy is dependent to some degree on the assumed functional form of the deterministic path of ρ_t, as in (10) for example. If the true path is nonlinear, then the estimates obtained from a linear model may be significantly distorted. This is true for a rolling windows approach as well, insofar as the linear model in (9) is always used in the second stage. However, by plotting the coefficient estimates from the rolling windows across time, one can observe if the shape of the evolution of the coefficient is clearly nonlinear, and then easily adjust the second-stage regression appropriately. Of course, one could always augment the models to include quadratic terms, for example, to capture nonlinearities, although this becomes more problematic when ρ_t does not make a smooth transition but rather makes a discrete jump due to a structural break. We find this to be an important consideration when interpreting our empirical results in the next section.

3.3 Results

In this section, we apply the time-trending regression model and rolling windows model outlined in the previous sections to our panel of RER data introduced in Section 2. We consider the application of both LSDV and DCCE estimators within each estimation framework. In the case of DCCE estimation, we utilize the excellent tool set developed by Ditzen (2018) and follow the guidance of Chudik and Pesaran (2015) by including a number of lags of the cross-sectional averages equal to the floor of $T^{1/3}$, where T represents either the time dimension of the full sample or the windows size, depending on the context. We focus our analyses on two groups of data: 1) a panel of 20 developed "OECD" countries, and 2) a fuller panel of 66 developed and developing countries.

In the following sections, the reported "half-life" results represent the time it would take in years to revert halfway from a shock *presuming that the autoregressive coefficient continues to change over future periods* as dictated by its fitted values. However, if the reversion to PPP requires more periods than the number of available future-period estimates, we simply repeat the use of the fitted value in the final available period. In the case of decreasing coefficients, this results in slightly smaller half-lives than if we assumed that the coefficient were to remain constant from the time of the shock over all future periods, as is assumed in (7). This predominantly affects early-period estimates, since the two methods of reporting half-lives will converge to the same values as we approach the time horizon.

3.3.1 Time-Trending Regression Model Results

Estimation results for the time-trending regression model are presented in Table 3. First, focusing on the DCCE estimator results, which control for common correlated effects as well as heterogeneous coefficients, we find coefficients that both started fairly high in value and exhibited negative changes that are statistically significant over the time horizon, for both groups of countries. The corresponding half-life estimates based on these coefficients are presented in Figure 2. For the OECD group of countries in particular, the estimated coefficient on $t \cdot q_{i,t-1}$ is highly significant and suggests $\hat{d\rho} = -0.0035$, which can be interpreted as a fall in the half-life to PPP convergence of about 3 weeks per year during the mid-point of the time sample.

The estimated coefficient is similarly significant and negative for the sample including all countries, but is smaller in magnitude. However, the estimated initial coefficient value at the beginning of the sample period is relatively low. This generates an interesting comparison between the groups in Figure 2, where we observe the OECD group of countries experiencing relatively higher half-lives, but a greater rate of decline, such that they "catch-up" to the group of all countries by the end of the sample period. By 2015, the results suggest both groups of countries exhibit similar half-lives of just below 2 years. Since the group of all 66 countries also contains the sub-sample of 20 OECD countries, the results seem to indicate that most of the movement in the larger group is emanating from that smaller group of advanced economies. This seems to suggest some sort of dichotomy in the ways that developed and developing economies have experienced PPP convergence. An interesting question then is whether this estimated difference in dynamics will continue to play out over future years, with advanced economies not being in the process of *converging* to a global average but rather *overtaking* other less-developed countries in terms of price flexibility? Moreover, if historical trends continue, these results may foreshadow important implications for the conduct of policy in OECD countries that experience less and less international price stickiness.

Interpreting the coefficients from the LSDV estimator is a bit more problematic. For OECD countries, the estimate is also negative but not significantly different from zero, whereas the estimated coefficient for all countries is positive. Why the discrepancy? Our preferred explanation is that the lack of accounting for common correlated effects, particularly the effect of the euro's introduction, is significantly coloring the results. In this case, the ability to visualize the data proves incredibly helpful, and so we next turn to the results from the rolling windows model to aid in our explanation. Though the DCCE estimator provides what we believe to be more rigorous and accurate results, we present the LSDV results here partially as a point of comparison to earlier studies and an explanation as for why earlier work may not have revealed much dynamic adjustment.

Time-Trending Regression Model Estimation

	OECD Countries		All Countries	
	LSDV	DCCE	LSDV	DCCE
$q_{i,t-1}$	0.9493***	0.8862***	0.8155***	0.8487***
	(0.0672)	(0.0372)	(0.0241)	(0.0473)
$t \cdot q_{i,t-1}$	−0.0016	−0.0035***	0.0010*	−0.0027*
	(0.0013)	(0.0012)	(0.0005)	(0.0014)
R^2	0.93	0.72	0.90	0.68
N	1100	1060	3630	3498

Table 3: Estimation results for full sample, 1960-2015, using time-trending regression model in (11). Standard errors in parentheses. *, **, and *** represent statistical significance at the 10, 5, and 1% levels, respectively. For brevity, estimates of country-specific covariates included in the regressions are not reported.

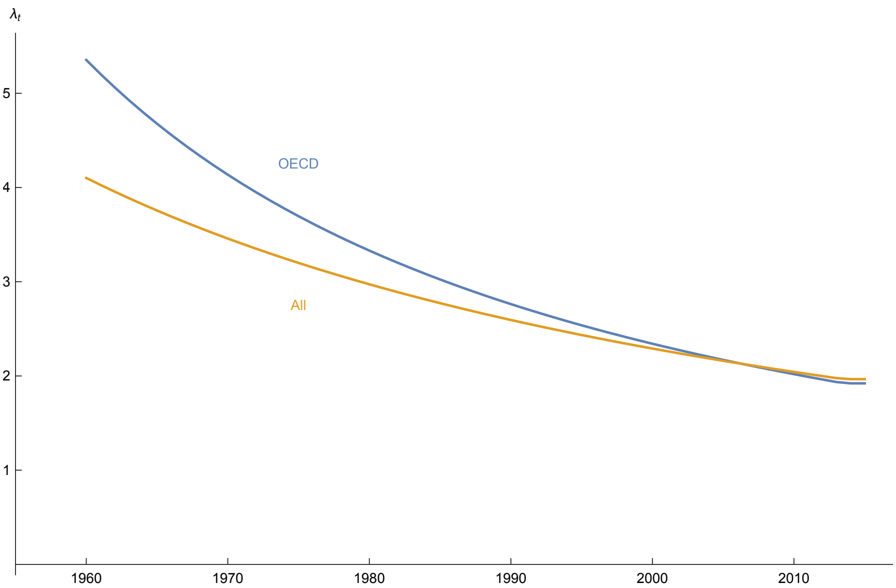


Figure 2: Estimated half-lives from time-trending regression model using DCCE estimator.

3.3.2 Rolling Windows Model Results

In this section, we graphically present the results from the estimation of the rolling windows model in 3.2.1. We again apply the model, using both LSDV and DCCE estimators on each window, to the OECD and full-country panels. Performing this analysis of course requires a somewhat arbitrary yet potentially important choice of window size, w, to be used. In our review of past literature, unfortunately, we couldn't find much theoretical guidance on the optimal number of periods to use in this context. The choice involves resolving an inherent tension between 1) improving the accuracy of the estimation *within* each window by increasing w, and 2) improving the accuracy of the estimation *between* windows by reducing w and thus increasing the number of observations available in the second-stage regression. Based on our own Monte Carlo analysis, we observed that a ratio of window size to full time dimension of $w/T \approx .5 - .7$ appears to minimize the bias in estimating $d\rho$. Therefore, we choose a window size of 30 years relative to the full 55 years of complete observations[10] available in our data set.

We first present results from the use of a standard LSDV estimator on the OECD panel.[11] The estimated autoregressive coefficients from fitting the rolling model in (8) are presented in Figure 3. Prior to the year 2000, the results pretty closely mirror our findings from the time-trending regression model with a DCCE estimator, with the coefficient starting off pretty high at just below 0.9, then falling substantially in subsequent years. However, the introduction of the euro starting in 1999 has a profound impact on the estimates. Out of the 20 countries in this OECD panel (not counting the US as the base country), 11 of them are members of the eurozone by the end of the sample in 2015, representing an important cohort with the ability to greatly influence the overall results. This can be seen most clearly just by observing the sudden appearance of an extremely large confidence interval associated with the first window to include the year 1999, the first year of the euro, due to the substantial shocks experienced by several members countries' RERs. The overall pattern in subsequent windows resembles a discrete jump to higher-level coefficients before a return a gradual decline. Altogether, the second-stage regression yields a dashed line-of-best-fit with a slightly positive slope. After viewing the data in this way, however, it seems apparent that the impact of euro adoption is driving the increase in half-lives that contradicts our earlier results using the DCCE estimator.

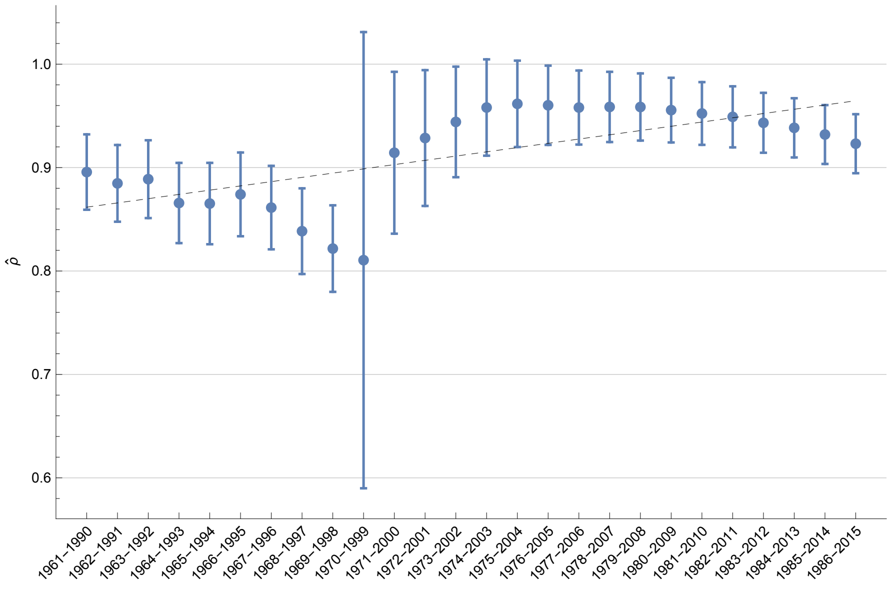


Figure 3: Rolling-window coefficient estimates for panel of OECD countries using LSDV estimator. Bars represent 95% confidence intervals. Dashed line-of-best-fit from second-stage regression of coefficients on windows.

We consider an adjustment to the estimation model to address this apparent discrete structural shift by augmenting the LSDV rolling regression model with an additional dummy variable:

$$q_{i,t} = \alpha_i + \rho_j \cdot q_{i,t-1} + \delta_j \cdot d_{i,t} + e_{i,t}, \quad t = j, \ldots, j + w - 1 \tag{12}$$

[10] We lose the initial year of 1960 because of the need for a lag.

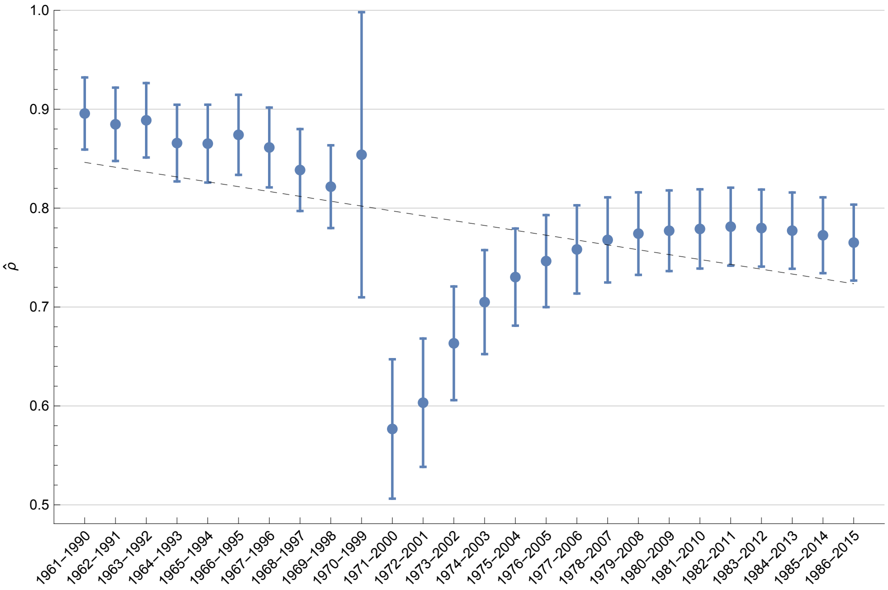
[11] We also implemented the Nickell and time-aggregation bias adjustment formula from Choi et al. (2006) in our rolling windows framework as a robustness check, and found that while the estimated levels of the coefficients were generally slightly higher, the overall shape of the analysis did not differ substantially.

where $d_{i,t}$ equals one in a year when that country is a member of the euro area, and zero otherwise. Note that the estimated coefficient is common to all countries, as introducing country-specific euro-impact dummy coefficients would introduce a burdensome number of parameters to be estimated with the limited number of observations available in each window. After applying this augmented model to the data, we see that the inclusion of the euro dummy effectively just shifts the post-1999 estimates downwards, especially in the early windows. Altogether, this creates a line-of-best-fit in the second-stage regression that now again exhibits a downward trajectory. Depending on one's interpretation of the euro dummy, the story is either that there is a more-or-less uninterrupted downward trend over the whole sample, or there is a one-time upward jump when the euro is introduced, after which the global economy recalibrates and returns to a downward trend. Our view is that the euro dummy is effectively making explicit (in a relatively ham-fisted way) the most important of many potential common correlated effects that are treated as unobservable by – but incorporated into – the DCCE estimator and reemphasizes the importance of its use.

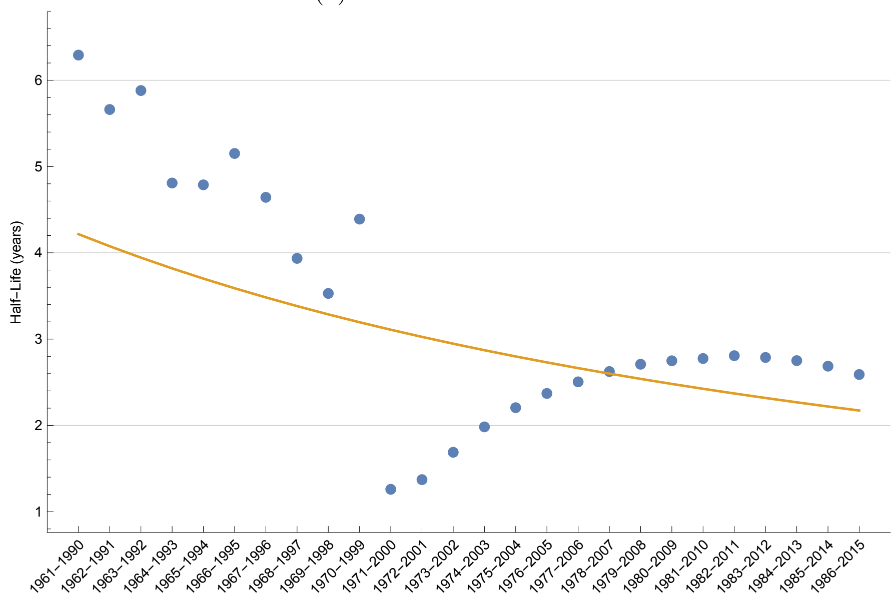
Before moving on to the DCCE results, it's worth noting that the introduction of the euro was not the only major exchange rate regime change to occur during the sample period. The collapse of the Bretton Woods international monetary system during the early windows of course involved most of the countries that comprise the OECD group in our data. An obvious question then is why the latter euro event shows up so clearly in the estimates, whereas the former does not? One explanation is that the timing of the euro introduction was rigidly scheduled to occur, "ready-or-not" (and several of the initial member countries were perhaps not ready, judging by the Maastricht criteria), whereas disengagement from Bretton Woods was somewhat drawn out and flexible. Although the "Nixon shock" was abrupt and unanticipated when it occurred in August 1971, it was followed by the Smithsonian Agreement in December and ongoing attempts to shore up the system for another 16 months or so until the EEC countries and Japan elected to drop their pegs to the US dollar in 1973.

More to the point, the biggest impact on the RER when dropping a peg would likely arise via the nominal exchange rate (presuming some degree of misalignment, which may not exist), but this would represent a transitory shock rather than something affecting the structural speed of mean-reversion. The abandonment of a currency peg, under which all domestic goods are still priced in local-currency units, is fundamentally different from joining a monetary union, whereby all domestic goods must be newly priced in a nascent currency. In theory, the new prices in the monetary union would simply be translations of the original prices using the official rate of exchange between the new currency and the one being phased out. In practice, however, the adjustment may not be so straightforward. This perhaps contributes to a shift in a country's long-run PPP convergence level, as opposed to simply representing a transitory RER shock, and therefore the regression model benefits from the flexibility to estimate a distinct euro-specific mean.

Next, we present in Figure 5 our preferred set of results derived using a DCCE estimator



(a) Coefficient Estimates



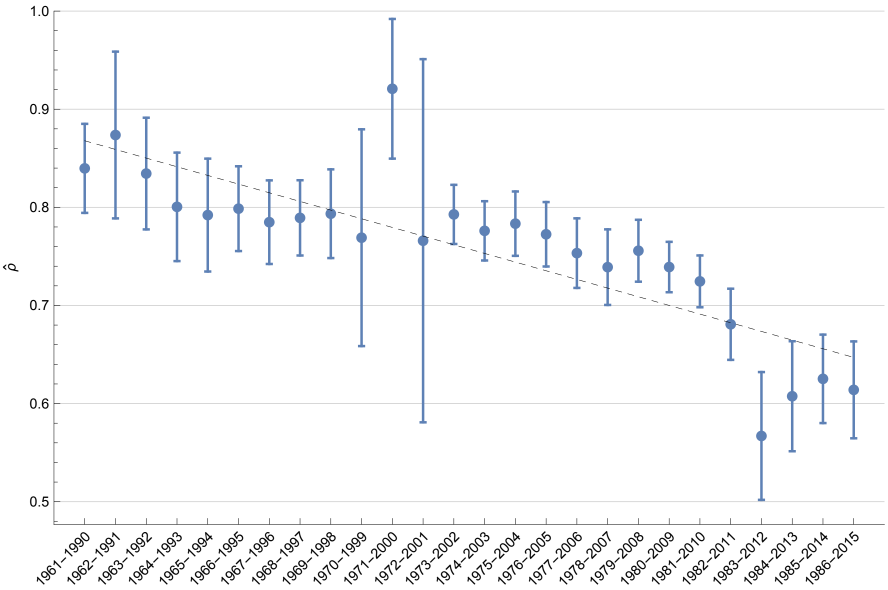
(b) Half-life Estimates

Figure 4: Rolling-window estimates for panel of OECD countries using LSDV estimator augmented with euro dummy. Bars represent 95% confidence intervals. Dashed line-of-best-fit from second-stage regression of coefficients on windows. Half-life point estimates from application of (7) to coefficient point estimates. Half-life line from simulation of forward-looking AR(1) processes using fitted coefficient values.

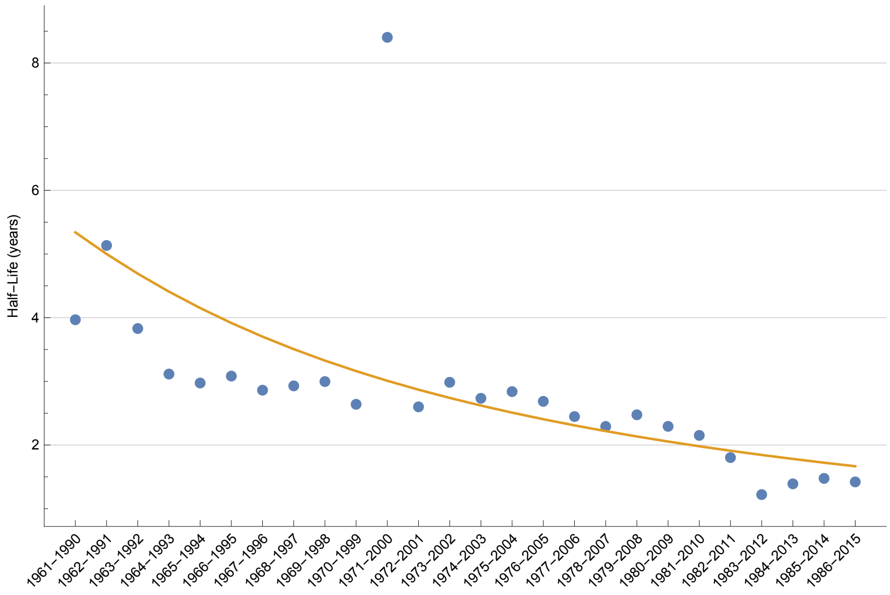
on the panel of OECD countries. We again observe some noise in the estimation of windows immediately after the introduction of the euro; however, despite the absence of a dedicated euro dummy in this context, the overall picture is that of a relatively clear downward trend in the estimated half-life of PPP convergence. Reassuringly, these results more or less tell the same story as our results from the time-trending regression model. The total evolution is rather striking: a fall in half-life over the sample period from a high at the beginning of 5.3 years to a low at the end of only 1.7 years! This represents a dramatic shift in the character of the global economy, which we attribute to the gradual forces of globalization and technology over these decades.

Finally, we present results from applying the DCCE estimator in a rolling windows model to the full panel of 66 countries in Figure 6. Extending the DCCE analysis to a broader set of economies does not substantially change the shape of the results, which exhibit a downward trend. There is an uptick in the estimates during some of the middle periods, but again this suggestively occurs in the windows corresponding to the early years after the introduction of the euro. The overall movement is from a high of 3.6 years to a low of 2.0 years. While not as dramatic as the dynamics within the OECD group, the results are again similar to the time-trending regression model results in the previous section and suggest perhaps some differences in dynamics among developed and developing economies.

Note that in all cases, the estimated dynamics have brought down the estimated half-life of PPP reversions to around 2 years by the end of the sample period in 2015, well below Rogoff's (1996) previous consensus of 3–5 years. Moreover, we find no evidence to suggest that this decreasing trend won't continue to extend into the future, further increasing the speed of international price convergence. However, if the linear trend of the estimated coefficients does continue, the rate of decline in the half-life will slow down because 1) the nonlinear transformation requires larger and larger decreases in the level of the coefficient to generate steady decreases in the corresponding half-lives, and because 2) it would eventually run up against a natural lower-bound, i.e. the speed of domestic price convergence.

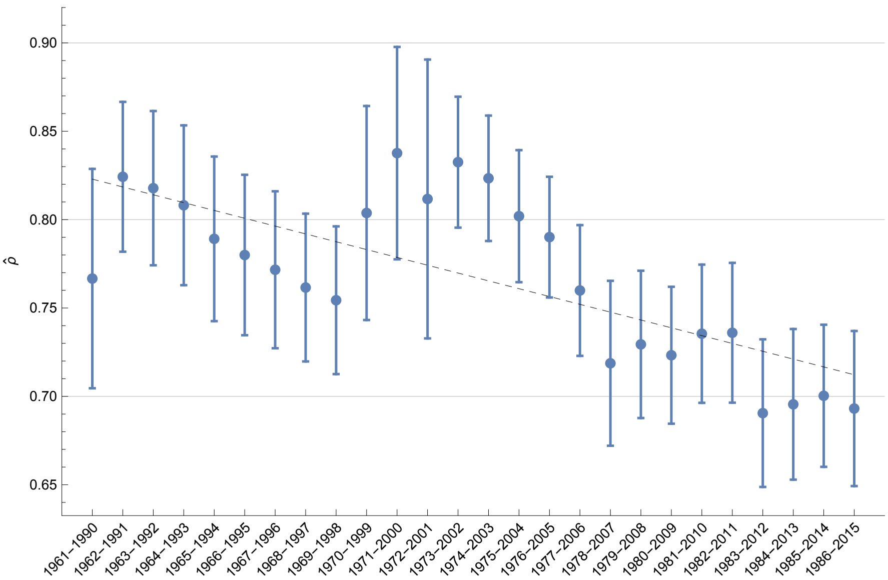


(a) Coefficient Estimates

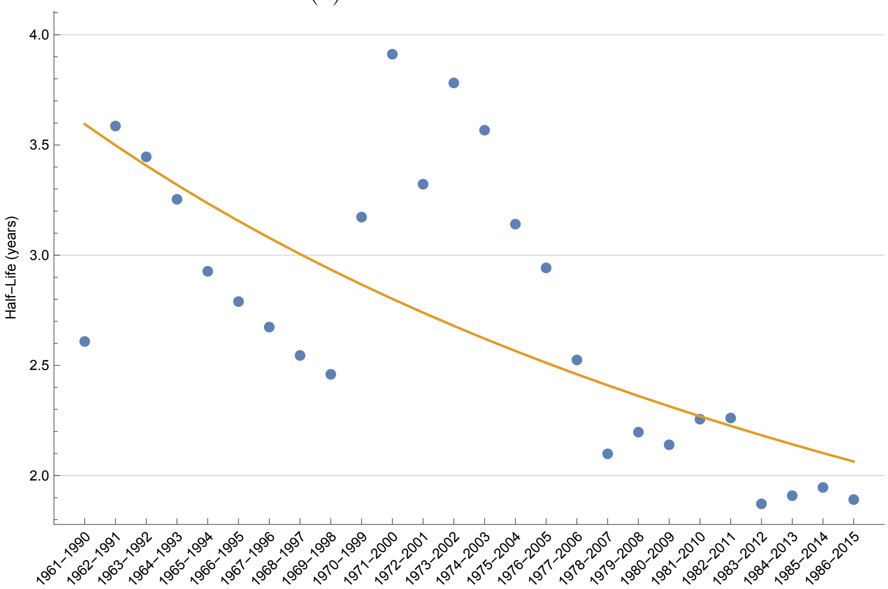


(b) Half-life Estimates

Figure 5: Rolling-window estimates for panel of OECD countries using DCCE estimator. Bars represent 95% confidence intervals. Dashed line-of-best-fit from second-stage regression of coefficients on windows. Half-life point estimates from application of (7) to coefficient point estimates. Half-life line from simulation of forward-looking AR(1) processes using fitted coefficient values.



(a) Coefficient Estimates



(b) Half-life Estimates

Figure 6: Rolling-window estimates for panel of all countries using DCCE estimator. Bars represent 95% confidence intervals. Dashed line-of-best-fit from second-stage regression of coefficients on windows. Half-life point estimates from application of (7) to coefficient point estimates. Half-life line from simulation of forward-looking AR(1) processes using fitted coefficient values.

4 Theoretical Explanations

Given the empirical results, how can we explain the increased speed of convergence to international PPP? In this section, we provide a simple theoretical explanation, in additional to empirical evidence, based on changes in the underlying ways in which real exchange rates and national consumer price indices (CPI) are constructed.

To begin, we return to the definition of a country's real exchange rate in (1), but now in level terms:

$$Q_t = S_t \cdot \frac{P^*}{P_t} \tag{13}$$

where Q_t is the real exchange rate in period t, S_t is the nominal exchange rate (in local currency units per foreign currency unit), and P_t is the domestic price level. For simplicity of exposition, we assume that the foreign price level, P^*, is constant over time and does not respond to shocks to the economies of either country.

This definition of the RER makes it clear that there are exactly three components that can possibly react to any shock. The question at hand is what is driving the *speed* at which these three variables change to bring the countries back to equilibrium. Note that this is different from considering the sources or sizes of the shocks themselves. For example, if both countries experience simultaneous symmetric shocks to their respective national price levels, then the changes in the numerator and denominator in (13) would cancel each other out, and there would be no change in the RER. While this might explain how a particular pair of countries are able to stay in alignment with each other relatively well over time, it does not reveal anything about how quickly the country pair moves back to PPP once they actually do experience a shock to their RER.

We now impose a theoretical framework on the definition in (13) by making two assumptions: 1) national prices are geometric averages of the domestic prices on tradable and nontradable goods, and 2) the prices on tradable and nontradable goods are updated by firms stochastically over time, *a la* Calvo (1983). The updated definition can now be expressed as

$$\begin{aligned} Q_t &= S_t \cdot \frac{\left(P^{T*}\right)^{\phi^*}\left(P^{N*}\right)^{(1-\phi^*)}}{\left(P_t^T\right)^{\phi}\left(P_t^N\right)^{(1-\phi)}} \\ &= S_t \cdot \frac{\left(P^{T*}\right)^{\phi^*}\left(P^{N*}\right)^{(1-\phi^*)}}{\left(\gamma^T \cdot P_{\text{LOOP}}^T + (1-\gamma^T)\cdot P_{t-1}^T\right)^{\phi}\left(\gamma^N \cdot P_{\text{PPP}}^N + (1-\gamma^N)\cdot P_{t-1}^N\right)^{1-\phi}} \end{aligned} \tag{14}$$

where ϕ and ϕ^* represent the relative weighting of tradable goods in the domestic and foreign national price indices, respectively, and γ_T and γ_N represent the share of firms in each period that update their prices on tradable and nontradable goods, respectively. The domestic price on tradable goods moves toward the price level demanded by the law of one price, i.e. $P_{\text{LOOP}}^T = S_t \cdot P^{T*}$, and P_{PPP}^N is the domestic price level for nontradable goods that must prevail in the long run for absolute PPP to hold, once $P_t^T = P_{\text{LOOP}}^T$. Previous theoretical work, e.g. Carvalho and Nechio (2011) and Candian (2019), has applied nominal price-stickiness to dynamic open economy models in a rigorous way in order to model RER behavior. Our interest here is not in using the model to precisely match estimated empirical moments, but to demonstrate in a simple way how changing the underlying parameters that define domestic pricing behavior can then generate changes in the speed of mean-reversion in the RER.

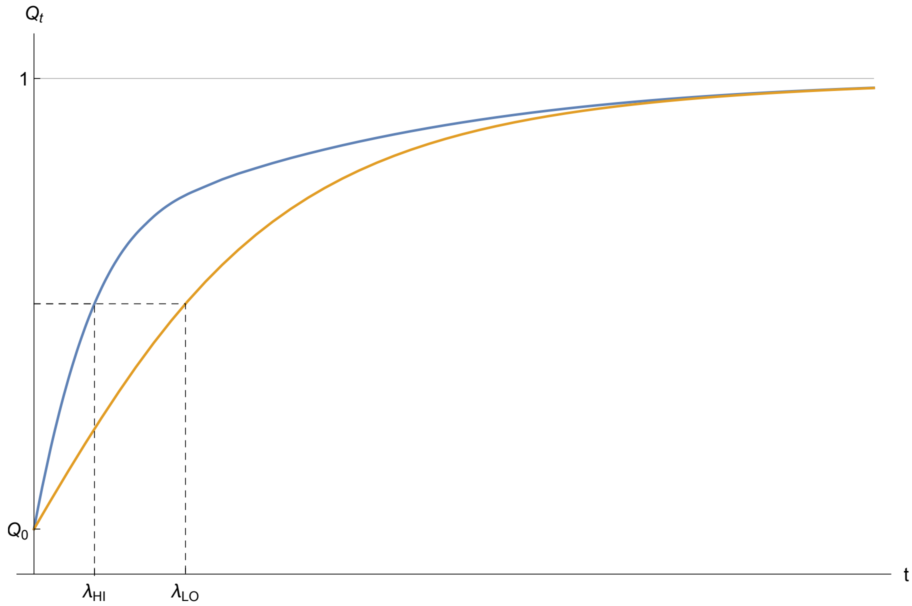


Figure 7: RER Convergence Speeds

Suppose that an economy is in a state of absolute PPP, such that $Q_t = 1$, before experiencing a negative shock to the nominal exchange rate in period $t = 0$ that permanently pushes the RER down to $Q_0 < 1$. Then, in accordance with (14), the economy's RER will slowly converge back to its PPP level as firms update domestic prices, based on the parameters of γ_T and γ_N. Now consider a baseline version of this economy ("LO") in comparison to an identical one with slightly higher values of γ_T and/or γ_N ("HI"). Because individual firms in the "HI" economy are updating their prices more quickly to align with the law of one price, its half-life of RER convergence, λ_{HI}, will be shorter than that of the "LO" economy, λ_{LO}, as demonstrated in Figure 7. The progress of technology in making access to market information easier and less costly than ever before, via mobile phones, the internet, etc., along with the greater interconnectedness of global markets certainly makes the story of stronger arbitrage forces raising the rate of Calvo-style price-updating over time seem plausible. Of course, as a point of theory, it's rather obvious that faster price-updating at the disaggregated level should generate faster movements in aggregate prices. Unfortunately, accurately estimating any potential changes in the degree of price-stickiness at the firm level is both challenging and outside the scope of this paper.

Rather, we propose a theoretical explanation for changes in RER convergence speeds based on the preceding logic but with a nuanced difference. First, we assume that firms operating in the tradable sector update their prices with greater frequency than those in the nontradable sector, i.e. $\gamma_T > \gamma_N$. This seems reasonable since international markets tend to be larger, more competitive, and more sophisticated relative to domestic ones, and thus more likely to experience arbitrage forces that engender price changes. Then, we propose that the source of the dynamic change generated in RER convergence is ultimately an increase in ϕ, i.e. the weight placed on tradable goods in constructing a nation's consumer price index. For an economy reacting to a negative nominal exchange rate shock, as before, the half-life of the response will again decrease from λ_{LO} to λ_{HI} as ϕ increases, generating RER paths similar to those in Figure 7. The mechanism driving this increased convergence speed is effectively the same as before, but the nuance lies in the idea that the fundamental frequencies of price-updating themselves are not presumed to be changing, but rather the categorization of the goods and services that the representative households are consuming. Therefore, if globalization brings about a shift in domestic preferences toward foreign goods, or lower transportation/trade costs allow for more previously nontradable goods to be effectively traded, or – more generally – a greater share of the domestic economy is subject to the influences of global markets, then we may well expect to see RERs react more quickly. The next section is devoted to examining in detail whether or not this shift toward tradability in national prices that are used to construct our measures of RERs has actually occurred over time.

5 CPI Tradability

In this section, we conduct an exercise aimed at understanding the source of the declining half-lives found in previous sections. In particular, we explore the extent to which the items that are included in our measure of prices, the CPI, have become more tradable or traded over the period of interest. From a theoretical standpoint, deviations from PPP arise when there is less trade and therefore less opportunity to arbitrage price differences across countries. An increase the number of goods and services traded or in the intensity of trade should in theory then be related to an increase in the speed of adjustment to shocks to PPP.

It is well known that world trade has increased dramatically over our sample period. However, it is not immediately obvious how this translates into the tradability of the components of the CPI. This is because the importance of various goods and services in the basket that comprises the CPI is changing over time and as the price of an item falls, it may constitute a smaller fraction to a consumer's average spending and, therefore, may have a smaller CPI weight. For example, the importance of food in the typical consumer's budget fell by about half from 1970 to 2015, caused primarily by a decrease in food prices. Because trade typically drives the price of a traded good down, increasing trade may in fact cause an item's CPI weight to decline. Therefore, the impact of increased trade on the CPI is ambiguous from a theoretical standpoint.

In what follows, we develop three different metrics to measure the extent to which the price index used in our empirical analysis has become more tradable or traded over time. The goal of this section is to explore a candidate explanation for the observed pattern of increased rate of PPP convergence.

5.1 Data and Methodology

We follow Johnson in constructing measures of the tradability of the CPI. We begin by compiling CPI weights provided by the Bureau of Labor Statistics (BLS) at five year intervals, starting in 1970 and ending in 2015.[12] It is important to note that the BLS re-weights the components of the CPI over time, to reflect the changing relative importance of certain goods and services in the average household's consumption basket. As such, we construct a time series of weights corresponding to each industry represented in the consumption basket. We then gather industry-level trade and output data from the Bureau of Economic Analysis's industry accounts. We use the annual Use Tables from these industry accounts in order to construct a measure of the fraction of the total final use that is either imported or exported in a given industry. Because these two data sources classify industries in different ways, it is necessary to create a concordance between the industry definitions in order to create our measure of the Tradable CPI. We again draw upon the work of Johnson[13] in order to concord the industry-level data across the two sources.

We then construct three different indexes relating to the tradability of the CPI. First, we use a crude definition of whether an industry is "tradable" by defining all goods as tradable and all services as non-tradable. Using this definition, we construct our first index by defining an indicator variable equal to one for goods producing industries and zero for service producing industries and multiplying this indicator by the industry-level CPI weights. Any changes in this index, then, reflect changes to the weights of the various industries, as the relative tradability of any given industry is being held fixed. To construct the second index, we follow the method in Johnson, creating an indicator variable that is equal to one if the fraction of final use in an industry that is traded (imported or exported) is greater than 15% and zero otherwise.[14] We then create a measure which we call the "Tradable CPI" by multiplying this indicator by the CPI weight for each industry. Lastly, we create what we call the "Traded CPI" by multiplying the fraction of final use in any industry that is traded, as measured as the sum of imports and exports in that industry divided by total final use, by the CPI weight for the industry. Both of the last two indexes then reflect both changes in relative importance of industries within the CPI and the amount of trade that occurs within each industry.

All of the constructed indexes gives us an idea of how many of the components of our measure

[12] Steve Reed and Jonathan Church were instrumental in gathering the CPI weights for years before 1987.

[13] We would like to thank Noah Johnson and Anya Stockburger at the BLS for providing us with the concordance that was constructed by Mr. Johnson. We had to modify his original concordance between the 2007 industry tables from the BEA to the 2007 CPI weights, as the 2007 industry tables include many more industries than the typical annual industrial input-output table.

[14] Following the existing literature, we set this threshold anywhere between 10% of final use to 20% of final use and there is little qualitative difference in our results.

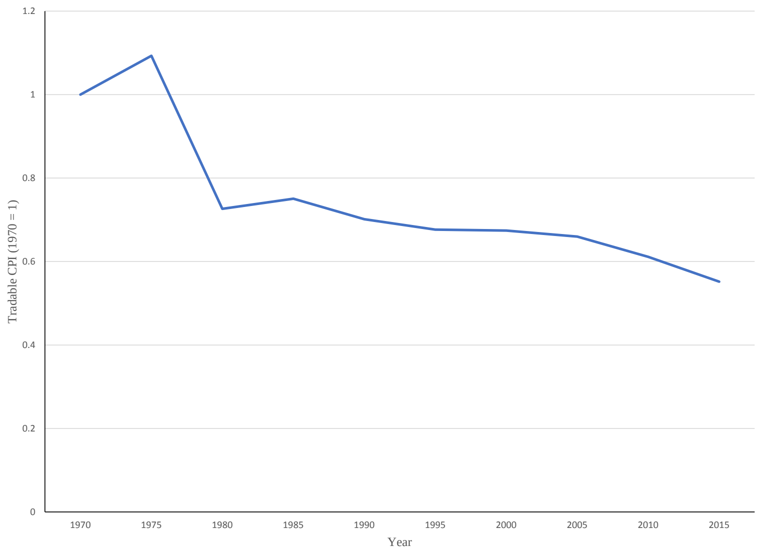


Figure 8: Tradability of CPI: Goods & Services

of prices should be subject to the law of one price. As is well known, if goods are not traded across countries, there is no reason to believe that the law of one price should hold. The goal of building these indexes is to get a sense of the extent to which the components of the CPI are either potentially or actually traded and, therefore, the degree to which the price index is influenced by trade.

5.2 Results and Discussion

Figure 8 displays the first of our measures, which is based on a crude definition of whereby goods producing industries are tradable and service producing industries are not. There is a substantial decline in this measure over time; in fact, using this crude measure, the tradability of the CPI falls by half during the 40 year period of interest. This is driven entirely by the decline in the relative importance of goods producing industries in the CPI over time, as U.S. consumers shift their spending more heavily towards services. We can conclude from this exercise that as the average consumer moves to consuming services, which are inherently more difficult to trade, it is not immediately obvious that the items in the CPI have become more traded/tradable over time. In fact, using this indicator, the decline in the adjustment speed to shocks away from PPP is somewhat puzzling. Using this measure, one would assume that the CPI has in fact become less tradable/traded over the period of interest, which in turn, would imply that we should expect the half-life of deviations from PPP to be increasing.

In contrast, Figures 9 and 10 show an increase in the tradability of the CPI, despite the fact that the basket weights placed on the most easily traded components, goods, are declining over the period of interest. The general pattern of increasing trade and tradability is evident in both figures. Even as the CPI weight of goods declines over time, both the number of industries which meet our criteria for being classified as "tradable" or "traded" and the relative importance of certain most-traded industries is increasing.

Figures 9 and 10 are instructive about the impacts of the extensive and intensive margins of trade, respectively, on the CPI. Consider first Figure 9, which reflects changes both in the CPI weights and in whether the various industries are "tradable," using dummy variables as defined above. Therefore, we can think of this index as representing the extensive margin for whether industries are "tradable." We see that there is a marked increase in this measure from 1970 to 2015. Figure 8 showed us that the CPI weights for many goods were decreasing over this time period, so the increase in this tradable CPI indicates that a number of industries switch from being "non-tradable" to "tradable" during this window. Therefore, from the increase in this index we can infer that a greater fraction of the items that comprise the CPI are becoming tradable over the period of interest. An increase in the tradabiliyt of the components of the CPI, in turn, indicates that to this price index is more likely to be subject to the law of one price.

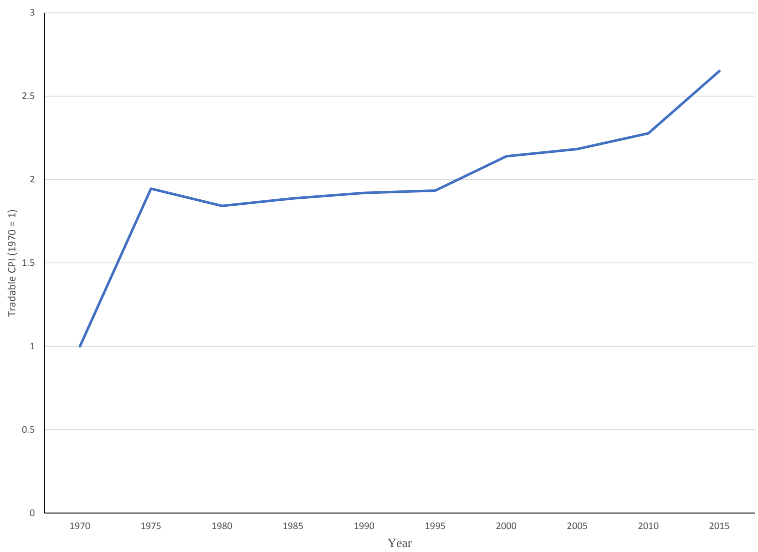


Figure 9: Tradability of CPI: 15% Threshold

Turning now to Figure 10, which combines the intensity of trade in each industry with the CPI weights, we can consider the intensive margin. Now, as we control for the fraction of final use that is traded in each industry, the traded CPI increases five fold from 1970 to 2015, indicating that the downward price effects of increasing trade are swamped by the increase in trade volume over this time period. An increase in the intensity of trade is an indication of the lowering of trade barriers, again suggesting that we should expect the components of the CPI to more closely adhere to the law of one price.

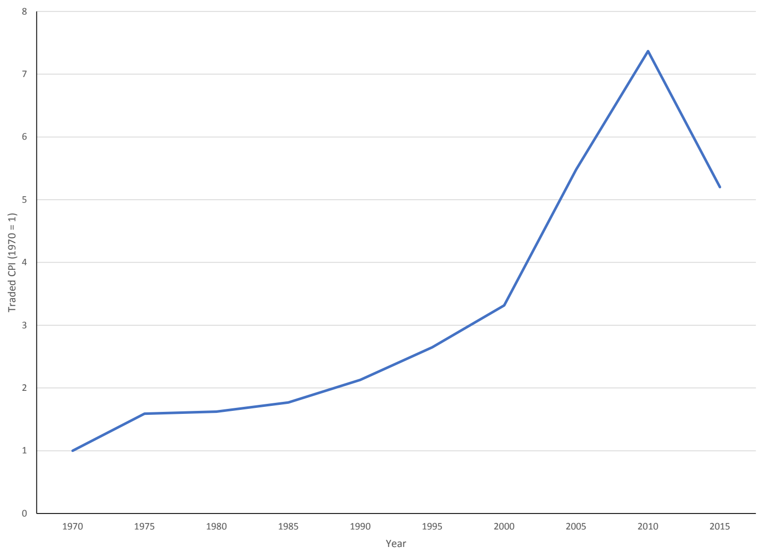


Figure 10: Tradability of CPI: Fraction Traded

Together, these results suggest that the decreasing half lives of deviations from PPP may be driven by an increase in trade. This increase in trade could have at its root cause many different factors, such as decreased tariffs, decreased transportation costs, or an increase in information on price differences across markets. Regardless of the source of increasing trade, its existence points to an increase in the role that trade and arbitrage play in determining the CPI. The increase in the tradability of CPI gives a potential explanation for our finding of increasing convergence speed back to PPP.

6 Conclusion

In the face of increased globalization and trade, a natural question arises as to what impact, if any, these forces have had on price differences across borders. In this paper, we aim to shed light on that question along a few dimensions. First, we build upon and contribute to the extant literature that estimates the speed of international price convergence and postulate that this convergence speed may be changing as the world economy becomes increasingly globalized. We depart from this literature by relaxing the assumption that of stability of the relationships underlying price movements across countries. Using an up-to-date econometric framework that controls for the sources of bias that have been identified in the literature, we estimate both the average speed of convergence and the *change* in this convergence speed for a large panel of countries from 1960 to 2015. When we estimate the average rate of PPP convergence, we find an estimate of 3 years, which is in line with what previous studies have shown. More interestingly, we find that this speed of convergence has fallen by 1.5 to 3 years over the sample period.

Our second contribution is showing the substantial sensitivity of these estimates to the selection of the time sample. Using Monte Carlo analysis, we show that when one estimates a dynamic coefficient process using a static-coefficient model, the resulting estimates are biased. Furthermore, the sign and magnitude of this bias depend crucially on the selection of the time horizon of the data. This result helps to contextualize differences in existing estimates of the convergence speed and the possibility that these differences are driven, in part, by the sensitivity of the estimates to selection of the sample time horizon.

Lastly, we conduct an analysis of the tradability of the price index to offer a potential explanation for the observed decline in the rate of PPP convergence. We show that, despite the fact that the most easily traded components of the CPI, goods, have declining CPI weights, the goods and services in the CPI become more tradable and more traded over the period of interest. We analyze the impact of both the extensive and the intensive margins of trade. To explore the extensive margin, we construct "tradable CPI" that combines CPI weights and a dummy measuring whether or not a item was traded in a given period and find that the increase in "tradability" resulted in a two and a half fold increase in the "tradable CPI" between 1970 and 2015. To analyze the intensive margin, we combine CPI weights with the fraction of industry-level final-use that is traded and find that by this constructed measure, the "traded CPI" has increased by roughly 500% over the same time horizon. These results suggest that a much larger fraction of the goods and services that make up the CPI basket are subject to international arbitrage forces at the end of the sample than the beginning, rationalizing the observed declines in convergence speed.

Our results provide insight for those seeking to understand and model international price co-movements. Furthermore, they point to several avenues for future research centered around the forces that could be driving changes in the PPP adjustment speed. We have explored one candidate explanation, but there are several others that could potentially be at play, including a decrease in financial frictions or increased availability of information across countries. These explanations potentially have different implications for how prices adjust across countries; therefore, understanding the source of the decrease in the speed of adjustment will serve to improve our theoretical understanding of international transmissions.

References

Abuaf, N. and Jorion, P. (1990). Purchasing power parity in the long run. *The Journal of Finance*, 45(1):157–174.

Astorga, P. (2012). Mean reversion in long-horizon real exchange rates: Evidence from latin america. *Journal of International Money and Finance*, 31(6):1529–1550.

Balli, H. O., Murray, C. J., and Papell, D. H. (2014). Median-unbiased estimation of structural change models: an application to real exchange rate persistence. *Applied Economics*, 46(27):3300–3311.

Bergin, P. R., Glick, R., and Wu, J.-L. (2017). conditional ppp and real exchange rate convergence in the euro area. *Journal of International Money and Finance*, 73:78–92.

Brown, R. L., Durbin, J., and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(2):149–163.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12(3):383–398.

Candian, G. (2019). Information frictions and real exchange rate dynamics. *Journal of International Economics*, 116:189–205.

Carvalho, C. and Nechio, F. (2011). Aggregation and the ppp puzzle in a sticky-price model. *American Economic Review*, 101(6):2391–2424.

Cheung, Y.-W., Chinn, M. D., and Pascual, A. G. (2005). Empirical exchange rate models of the nineties: Are any fit to survive? *Journal of international money and finance*, 24(7):1150–1175.

Cheung, Y.-W. and Lai, K. S. (2000). On the purchasing power parity puzzle. *Journal of international Economics*, 52(2):321–330.

Chinn, M. D. and Ito, H. (2006). What matters for financial development? capital controls, institutions, and interactions. *Journal of development economics*, 81(1):163–192.

Choi, C.-Y., Mark, N. C., and Sul, D. (2006). Unbiased Estimation of the Half-Life to PPP Convergence in Panel Data. *Journal of Money, Credit and Banking*, 38(4):921–938.

Chudik, A. and Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*, 188(2):393–420.

Chudik, A., Pesaran, M. H., and Yang, J.-C. (2018). Half-panel jackknife fixed-effects estimation of linear panels with weakly exogenous regressors. *Journal of Applied Econometrics*, 33(6):816–836.

Dhaene, G. and Jochmans, K. (2015). Split-panel jackknife estimation of fixed-effect models. *The Review of Economic Studies*, 82(3):991–1030.

Ditzen, J. (2018). Estimating dynamic common-correlated effects in stata. *The Stata Journal*, 18(3):585–617.

Edison, H. J. (1987). Purchasing power parity in the long run: A test of the dollar/pound exchange rate (1890-1978). *Journal of Money, Credit and Banking*, 19(3):376–387.

Frankel, J. A. (1986). *International Capital Mobility and Crowding Out in the U.S. Economy: Imperfect Integration of Financial Markets or of Goods Markets?* Federal Reserve Bank of St. Louis.

Frankel, J. A. and Rose, A. K. (1996). A panel project on purchasing power parity: mean reversion within and between countries. *Journal of International Economics*, 40(1-2):209–224.

Froot, K. A. and Rogoff, K. (1995). Perspectives on ppp and long-run real exchange rates. *Handbook of international economics*, 3:1647–1688.

Fung, H.-G. and Lo, W.-C. (1992). Deviations from purchasing power parity. *Financial Review*, 27(4):553–570.

Hegwood, N. D. and Papell, D. H. (1998). Quasi purchasing power parity. *International Journal of Finance & Economics*, 3(4):279–289.

Holmes, M. J., Otero, J., and Panagiotidis, T. (2012). Ppp in oecd countries: an analysis of real exchange rate stationarity, cross-sectional dependency and structural breaks. *Open Economies Review*, 23(5):767–783.

Imbs, J., Mumtaz, H., Ravn, M. O., and Rey, H. (2005). Ppp strikes back: Aggregation and the real exchange rate. *The Quarterly Journal of Economics*, 120(1):1–43.

Johnson, D. R. (1990). Co-integration, error and purchasing power parity between canada and the united states. *Canadian Journal of Economics*, pages 839–855.

Johnson, N. N. (2017). Tradable and nontradable inflation indexes: replicating new zealands tradable indexes with bls cpi data. *Monthly Labor Review, U.S. Bureau of Labor Statistics*.

Lothian, J. R. and Taylor, M. P. (1996). Real exchange rate behavior: the recent float from the perspective of the past two centuries. *Journal of political economy*, 104(3):488–509.

Manzur, M. (1990). An international comparison of prices and exchange rates: a new test of purchasing power parity. *Journal of International Money and Finance*, 9(1):75–91.

Manzur, M. and Chan, F. (2010). Exchange rate volatility and purchasing power parity: does euro make any difference? *The International Journal of Banking and Finance*, 7(1):104–123.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426.

Papell, D. H. (2006). The panel purchasing power parity puzzle. *Journal of Money, Credit, and Banking*, 38(2):447–467.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of applied econometrics*, 22(2):265–312.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. *Econometric Reviews*, 34(6-10):1089–1117.

Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of econometrics*, 68(1):79–113.

Phillips, P. C. and Sul, D. (2007). Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. *Journal of Econometrics*, 137(1):162–188.

Rogoff, K. (1996). The purchasing power parity puzzle. *Journal of Economic literature*, 34(2):647–668.

Taylor, A. M. (2001). Potential pitfalls for the purchasing-power-parity puzzle? sampling and specification biases in mean-reversion tests of the law of one price. *Econometrica*, 69(2):473–498.

Taylor, A. M. (2002). A century of purchasing-power parity. *Review of economics and statistics*, 84(1):139–150.

Vigfusson, R. J., Sheets, N., and Gagnon, J. (2009). Exchange rate passthrough to export prices: Assessing cross-country evidence. *Review of International Economics*, 17(1):17–33.