An Econometric Approach to the Construction of Complete Panels of Purchasing Power Parities: Analytical Properties and Empirical Results

D.S. Prasada Rao*, Alicia N. Rambaldi, Howard E. Doran

March 17, 2011

School of Economics, The University of Queensland. St Lucia, QLD 4072. Australia

Abstract

The Penn World Tables (PWT) provide estimates of purchasing power parities (PPPs) for 180 countries and fifty years. Despite their popularity, the PWT are anchored on a single *benchmark comparison* from the International Comparison Program conducted every five to ten years. An econometric approach which utilizes PPP data from all the ICP benchmarks and national growth rates is proposed and its analytical properties are derived. A state-space representation and a smoothing algorithm are used in generating a complete panel of PPPs with standard errors. PPPs for 141 countries covering 1970-2005 generated using the new method are compared to PWT 6.3.
JEL: C53, C33.
Keywords: Purchasing Power parities, Penn World Tables, State-space models, Spatial autocorrelation, Kalman Filter

Cross-country comparisons of growth and economic performance require economic aggregates such as the gross domestic product, consumption and investment in real terms and expressed in a common currency unit. Purchasing power parities (PPPs)[1] reflecting the relative price levels in different countries are considered superior to the market exchange rates for this

*Corresponding author. School of Economics, The Universtiy of Queensland, St Cucia, QLD 4072, Australia. Email: p.rao@economics.uq.edu.au

[1]PPP of a currency is defined as the number of currency units required to buy the same goods and services that can be bought with one unit of reference/numeraire currency. For example, a PPP of 130 Japanese yen per US dollar means that we need 130 yen to purchase the same goods and services that can be bought with one US dollar. The Big Mac index is a good example of a PPP which is based only on a single product.

purpose[2]. Catch-up and convergence of real incomes has been a subject of intense research among development economists (Pritchett, 1997; Quah, 1997; Durlauf et al., 2005; Barro and Sala-i Martin, 2004) – these analyses typically rely on real per capita income data derived using PPPs. The recent debates on globalization and inequality are anchored on PPP-converted income data (Theil, 1989; Milanovic, 2002; Sala-i Martin, 2006; Chotikapanich et al., 2007. The Human Development Index (UNDP, various) and estimates of regional and global incidence showing the number of people under \$1/day and \$2/day regularly published in WDI (various years) rely on PPP data.

The diverse range of applications listed above as well as researchers, analysts and policy makers at the national and international levels critically depend upon the availability of reliable PPP data for a large number of countries covering a long period of time. The reality is that PPP data are sparse and are available for only selected benchmark years from the International Comparison Program (ICP)[3]. The Penn World Tables pioneered by Summers and Heston [1988, 1991] providing PPPs and real incomes covering in excess of 180 countries and a 50-year period largely filled the gap[4]. Consequently, PWT has become one of the most widely used and cited source of data in economics. PWT 6.3 was recently released covering up to the year 2005. Details of the construction of PWT are available from Summers and Heston [1991] and Heston et al. [2006].

Despite the popularity enjoyed by the PWT, there are several directions in which the methodology that underpins the PWT can be improved. First, the PWT construction is anchored on a single benchmark of the ICP. The recent versions of PWT (versions 6.1, 6.2 and 6.3) are all based on the 1996 PPP data from the ICP. Second, the PWT uses a two-stage procedure. In the first stage, PPP benchmark data are used in a regression model to explain national price levels and the estimated model is used in extrapolating PPPs for countries that have not participated in the benchmark year on which the PWT is anchored. In the second stage, the PPPs for participating countries and the extrapolations to non-participating countries are extrapolated backwards and forwards in time using national price deflators and growth rates. Given the non-econometric nature of the extrapolations,

[2]Issues relating to the use of PPPs versus market exchange rates are discussed in detail in Kravis and Lipsey, 1983, in the final report of the 2005 ICP available on the World Bank `http://siteresources.worldbank.org/ICPINT/Resources/ICP_final-results.pdf` and the final report of ICP Asia-Pacific (Asian Development Bank, 2007).

[3]The ICP began as a research project at the University of Pennsylvania conducted by Kravis, Summers and Heston. Currently ICP has grown into a major international statistical program conducted by the World Bank under the guidance of the UN Statistical Commission. The first benchmark comparison was in 1970 and the most recent being the 2005 round of the ICP.

[4]Maddison [1995, 2007] provides real income series for a large number countries covering a long period of time since 1820 expressed in 1990 international dollars.

the PWT extrapolations have no measures of reliability[5], for example in the form of standard errors, associated with them.

The main objective of this paper is to propose a new approach designed to take the construction of PWT to the next generation of econometric sophistication and at the same time address some of the fundamental deficiencies of the PWT. The new methodology proposed here combines PPP data available from all the benchmarks of the ICP, instead of relying on a single ICP benchmark, with data on (relative) price movements available from the national accounts in the form of implicit deflators. The regression methodology used in the PWT is further strengthened with the specification of spatially autocorrelated disturbances in the regression model of the national price levels. Finally, the proposed econometric method also generates standard errors of the associated PPPs.

The current paper describes the complete version of the method in contrast to that described in * et al. [2010], referred to as RRD from here on. The RRD version is designed to disseminate a preliminary version of the work and it mainly focuses on the details of the regression model used in explaining national price levels[6] and on a special case of the general methodology described here.

The structure of the paper is as follows. Section 1 establishes the relevant notation and provides a statement of the problem. In section 2 we establish the basic framework of the method. Section 3 is devoted to the econometric formulation of the model including a discussion of the stochastic assumptions. The econometric model is given a state-space representation which makes it possible for us to use Kalman filtering and smoothing methods to generate optimal predictions. The analytical properties of the methodology proposed are discussed in section 4. Proofs of the main propositions are provided in the Appendix. In particular, we are able to show that the methodology proposed here is invariant to the choice of the reference country and that the model is flexible enough to either track the PPPs in different benchmarks or to track the national price movements[7]. Details of estimation of the state-space system are provided in section 5. Empirical results illustrating the feasibility and versatility of the proposed methodology are presented in section 6. The basic data used in the study and their sources are described in detail. We present predicted panel of PPPs for 141 countries covering the period 1970 to 2005. As the year 2005 is the latest benchmark year for the ICP, we present predicted panel of PPPs with and without the use of data

[5]PWT, however, provide some indication of the reliability as perceived by the compilers and are usually expressed in the form a quality rating in the range A to D.

[6]We avoid duplicating this material by simply referring to the corresponding sections of RRD.

[7]Due to inherent inconsistencies between benchmarks and national price deflators (and hence the national growth rates), it is not possible for the methodology to generate a series that is capable of tracking both at the same time.

from the 2005 benchmark. Predicted PPP series for a set of selected countries are further analyzed and compared with the corresponding PWT 6.3 series. The paper concludes with a few remarks in section 7.

1 Notation and a Statement of the Problem

We consider the general case with N countries (i=1,2,...,N) and T time periods (t=1,2,...T). Let PPP_{it} and ER_{it}, respectively denote the purchasing power parity and exchange rate of the currency of the *ith* country expressed in terms of the currency units of a reference country. PPPs reflect the general price levels whereas the exchange rates reflect the value of the respective currencies and generally reflect the demand for the currencies involved. Based on these, we define the *national price level* for country i in period t, as:

$$R_{it} = \frac{PPP_{it}}{ER_{it}} \tag{1}$$

where R_{it} is the national price level for country i in period t. A value of the price level above 1 indicates that prices are higher relative to the reference country and vice versa.

The Problem

Given the importance attached to PPPs in converting national economic aggregates into a common currency unit, the ideal situation is where PPP_{it}s are observed for every country in all time periods. However, the reality is quite different. The main source of these PPPs is the International Comparison Program which oversees the collection of data and compilation of PPPs for all the countries participating in the Program. Given the resource intensive nature of the underlying price collections, participation in the ICP has been limited and, consequently, the matrix of available PPPs is quite sparse. Table 1 shows the availability of PPP data from different rounds of the ICP.

[Table 1 here]

In practice we wish to have PPPs for all the countries spanning a long period of time. For example, if we have 180 countries and a period covering 1970 to 2005, then it is easy to see that the information available from the ICP and the OECD-Eurostat sources shown in Table 1 would sparsely cover the space-time tableau of PPPs of dimension 180 × 35. So our main problem is to develop an econometric technique that allows us to generate optimal predictions for the missing PPP data in the tableau. A simple approach to this problem would be to take PPPs from a single benchmark, say the 1985 benchmark and extrapolate the PPPs of the 64 participating countries to the 180 countries we wish to cover and, then, extrapolate over time using relative price movements. This is the approach that underpins the PWT (Heston et al., 2006) and the Maddison, 1995, 2007 series. However, the problem with this simple approach is that different benchmarks will produce different extrapolations and using a single benchmark amounts to discarding valuable information from the remaining benchmarks. The approach we propose here addresses this problem and uses information from all the benchmarks.

2 The Basic Framework

The basic building blocks that underpin the econometric methodology are described here. This description draws from a more elaborate exposition of the material in RRD. In this paper we consider the most general econometric model of which the case discussed in RRD would be a special case. We highlight the significant differences when we discuss the econometric model and its state-space representation.

(i) We begin with PPP data from the ICP and postulate that the observed PPPs are true values contaminated with noise and measurement error. Let $p_{it} = \ln(PPP_{it})$ be the logarithm of the true PPPs. The observed PPPs from the ICP are related to the true PPPs through the following equation:

$$\tilde{p}_{it} = p_{it} + \xi_{it} \tag{2}$$

where $\tilde{p}_{it}$ is the ICP benchmark observation for participating country i at time t; ξ_{it} is a random error accounting for measurement error with $E(\xi_{it}) = 0$[8]

(ii) We take into account the fact that the numerical value of the PPP for the reference/numeraire country is set at 1, we have the condition in logarithms as:

$$p_{1,t} = 0, \qquad t = 1, 2, \ldots T \tag{3}$$

where country 1 is taken to be the reference country without loss of generality.

(iii) The most important element in the extrapolation strategy is the theoretical model that provides a link between national price levels and a range of observable socio-economic variables within the countries. Drawing on the vast literature in this area (Kravis and Lipsey, 1983; Clague, 1988 and Bergstrand, 1991, 1996), we postulate a log-linear relationship

[8] We discuss the distributional assumption of disturbances in Section 4.

between the national price levels and a set of control variables[9]. We postulate the following model:

$$r_{it} = \beta_{0t} + \boldsymbol{x}'_{it}\boldsymbol{\beta}_s + u_{it} \quad \text{for all}\, i = 1, 2, ..., N \,\text{and}\, t = 1, 2, ..., T \tag{4}$$

where $r_{it} = ln(PPP_{it}/ER_{it})$; $\boldsymbol{x}'_{it}$ is a set of conditioning variables; β_{0t} intercept parameters; $\boldsymbol{\beta}_s$ a vector of slope parameters; u_{it} a random disturbance with specific distributional characteristics.

If estimates of β_{0t} and $\boldsymbol{\beta}_s$ are available, model (4) can provide a prediction of the variable of interest consistent with price level theory.

$$\hat{p}_{it} = \hat{\beta}_{0t} + \boldsymbol{x}'_{it}\hat{\boldsymbol{\beta}}_s + ln(ER_{it}) \tag{5}$$

where $\hat{p}_{it}$ is a prediction; $\hat{\beta}_{0t}$ and $\hat{\boldsymbol{\beta}}_s$ are estimates, and $\hat{p}_{it}$ is a prediction of $ln(PPP_{it})$.

In this paper we make use of (5) to generate predictions for all the cells (all countries in all years) in the panel of PPPs. This is a more general approach than that used in PWT or in RRD. In RRD, regression predictions are used for extrapolating PPPs only in benchmark years and that too only for the non-benchmark countries. This approach has implications to the specifications of the "observation equations" used in the state-space formulation in Section 3.2. The estimation of $\hat{\beta}_{0t}$ and $\hat{\boldsymbol{\beta}}_s$ is discussed in Section 5.

(iv) The main source of information used in the extrapolation of PPPs over time relates to the implicit price deflators data from the national accounts. The PPP of the currency of the ith country in period t, relative to the reference country, the US for example, can be updated to period $t+1$ by adjusting PPP_{it} for movements in the GDP deflators in country i and in the United States. Thus we have,

$$PPP_{i,t} = PPP_{i,t-1} \times \frac{GDPDef_{i,[t-1,t]}}{GDPDef_{US,[t-1,t]}} \tag{6}$$

Equation (6) defines the growth rate of PPP_{it}. We assume that the relationship (6) holds with a random disturbance term. The model in terms of $p_{it} = \ln(PPP_{it})$ is given by

$$p_{it} = p_{i,t-1} + c_{it} + \eta_{it} \tag{7}$$

where,
$c_{it} = ln\left(\frac{GDPDef_{i,[t-1,t]}}{GDPDef_{US,[t-1,t]}}\right)$; η_{it} is a random error accounting for measurement error in the growth rates

Equation (6) can be used in updating PPPs for any given year (especially a benchmark year) to all the years covered in the panel. Thus equation (6) can be independently used as a tool to fill the missing PPPs in the time-space panel of PPPs and it is used in the PWT and

[9] We discuss in detail the regression model specification and selection of variables in RRD

Maddison series. equation (7) plays a key role in the state-space representation of our model as it provides the "transition" equation for the model (see Section 3.2 for details).

3 Econometric Formulation and State-Space Representation of the Model

The objective is to produce a panel of predictions of p_{it} which optimally combines the information arising out of the four sources (i) to (iv) outlined in Section 2 above. The regression model explaining national price levels in equation (4) is incorporated into the following econometric model which is subsequently expressed in state-space form.

3.1 The Stochastic Assumptions

1. We start with the model used to explain the national price levels presented in equation (4). The errors u_{it} in (4) are assumed to be spatially correlated. The error follows a Spatial Error Model (see for example Chapter 2 of LeSage and Pace, 2009)

$$\mathbf{u}_t = \phi \mathbf{W}_t \mathbf{u}_t + \mathbf{e}_t \tag{8}$$

where $|\phi| < 1$ and $\mathbf{W}_t$ $(N \times N)$ is a spatial weights matrix. That is, its rows add up to one and the diagonal elements are zero. The term spatial in the present context refers to socio-economic distance rather than the traditional geographical distance[10]. It follows that $E(\mathbf{u}_t \mathbf{u}_t')$ is proportional to $\mathbf{\Omega}_t$, where $\mathbf{\Omega}_t = (\mathbf{I}-\phi\mathbf{W}_t)^{-1} (\mathbf{I}-\phi\mathbf{W}_t)^{-1'}$.

2. The measurement errors in the observation of $\ln(PPP_{it})$ during benchmark years, equation (2), are assumed spatially uncorrelated, but might be heteroskedastic. Thus, if ξ_{it} is a measurement error associated with the PPP of country i at time t, then

$$E(\xi_{it}) = 0; \quad E(\xi_{it}^2) = \sigma_\xi^2 V_{it} \tag{9}$$

where σ_ξ^2 is a constant of proportionality and $\boldsymbol{V}_t$ is defined below.

3. The measurement error in the growth rates, c_{it}, are assumed spatially uncorrelated, but might be heteroskedastic. Thus, η_{it} in (7) is assumed to have:

[10] In the empirical section we test for cross-sectional dependence and specify a model of socio-economic distance to obtain the weights.

$$E(\eta_{it}) = 0; \quad E\left(\eta_{it}^2\right) = \sigma_\eta^2 V_{it} \tag{10}$$

where σ_η^2 is a constant of proportionality and $\boldsymbol{V}_t$ is defined below.

4. The measurement error variance-covariance is of the form

$$\boldsymbol{V}_t = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \sigma_{1t}^2 \boldsymbol{j}\boldsymbol{j}' + diag(\sigma_{2t}^2, ..., \sigma_{Nt}^2) \end{bmatrix} \tag{11}$$

where, $\mathbf{j}$ is a vector of 1's, σ_{it}^2 is the variance of country i at time t, which we measure as the inverse of the a country's degree of development, and σ_{1t}^2 is the variance of the reference country.

This form of the covariance was derived from the definition of PPP (see RRD for detailed derivation) and it is sufficient for the invariance of the method to the choice of reference country (see Section4.5 and Appendix A for details on the invariance of the method). In the empirical implementation we model σ_{it}^2 as inversely related to GDP_{it}, the nominal per capita measured in \$US (exchange rates adjusted)[11]. This means that reliability of an observed PPP or growth rate is lower for low-income countries.

3.2 The State Space Representation

The econometric problem is one of signal extraction. That is, we need to combine all sources of "noisy" information and extract the signal from the noise. A state-space (SS) is a suitable representation for this type of problems. We start by writing equation (7) for the N countries to define the 'transition equation' of the SS:

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \boldsymbol{\eta}_t \tag{12}$$

where $\boldsymbol{\alpha}_t$ is the $N \times 1$ vector of unknown $\ln(PPP_{it})$ and it is the state vector in this representation; $\boldsymbol{c}_t$ is the observed growth rate of $\boldsymbol{\alpha}_t$ (this follows from equation (7) in Section 2); $\boldsymbol{\eta}_t$ is an error with $E(\boldsymbol{\eta}_t) = 0$ and $E\left(\boldsymbol{\eta}_t \boldsymbol{\eta}_t'\right) \equiv \mathbf{Q}_t = \sigma_\eta^2 \mathbf{V}_t$.

Equation (12) simply updates period $t-1$ $PPPs$ using the observed relative price changes, represented by $\boldsymbol{c}_t$,over the period .

[11]we use exchange-rate converted per capital income instead of PPP-adjusted per capita income mainly to avoid possible problems of endogeneity

As previously discussed, noisy observations of $\boldsymbol{\alpha}_t$ are given by: , a prediction $\hat{\mathbf{p}}_t$from the regression model (5),; a measurement $\tilde{\mathbf{p}}_t$by the ICP (2),, and by the condition in 3. Equation (5) relates the conditioning variables, $\boldsymbol{X}_t$, to the prediction $\hat{\mathbf{p}}_t$. Since the form of the observation equation of a SS model relates the observations (p_{1t}, $\hat{\mathbf{p}}_t$, $\tilde{\mathbf{p}}_t$) to the state vector $\boldsymbol{\alpha}_t$, it is convenient to re-write equation (4). Subtracting equation (5) from (1) we obtain:

$$\hat{p}_{it} = \alpha_{it} + (\hat{\beta}_{0t} - \beta_{0t}) + \boldsymbol{x}'_{it}(\hat{\beta}_s - \beta_s) - u_{it} \tag{13}$$

where α_{it} is an element of $\boldsymbol{\alpha}_t$.

Throughout the paper we will reserve the symbol $\boldsymbol{\theta}$ to represent the error in a current estimate of a parameter $\boldsymbol{\beta}$. Thus,

$$\hat{\theta}_{0t} = \hat{\beta}_{0t} - \beta_{0t} \text{ and } \hat{\theta}_s = \hat{\beta}_s \text{-} \beta_s \tag{14}$$

It is then possible to write equation (13) in the form

$$\hat{\mathbf{p}}_t = \boldsymbol{\alpha}_t + \mathbf{X}_t\boldsymbol{\theta} + \mathbf{v}_t \tag{15}$$

where $\boldsymbol{\theta} = [(\theta_{01}, ..., \theta_{0T})', \theta'_s]'$; $v_{it} = -u_{it}$; $\mathbf{X}_t$ is an $N \times K$ matrix with columns including time dummy variables and socio-economic variables.

Finally, in order to express these different types of observations (viz, those given by (2) and (15)) as a single equation, it is convenient to define the following 'selection matrices',

$\mathbf{S}_1 = [1, \mathbf{0}'_{N-1}]$ (selects the reference country $i = 1$)[12]; $\mathbf{S}_p$ is a known matrix which selects the N_t participating countries (excluding the reference country) in the benchmark year t; $\mathbf{S}_{np}$ is a known matrix which selects $(N - 1 - N_t)$ non-participating countries in the benchmark year t[13].

We are now able to consolidate these sources of information into a single equation on an 'observation vector' $\mathbf{y}_t$, viz

$$\mathbf{y}_t = \mathbf{Z}_t\boldsymbol{\alpha}_t + \mathbf{B}_t\mathbf{X}_t\boldsymbol{\theta} + \boldsymbol{\zeta}_t \tag{16}$$

with variables defined as follows:

i) Non-benchmark years:

$$\mathbf{y}_t = \begin{bmatrix} 0 \\ \mathbf{S}_{np}\hat{\mathbf{p}}_t \end{bmatrix}, \mathbf{Z}_t = \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_{np} \end{bmatrix}, \mathbf{B}_t = \begin{bmatrix} 0 \\ \mathbf{S}_{np} \end{bmatrix}, \boldsymbol{\zeta}_t = \begin{bmatrix} 0 \\ \mathbf{S}_{np}\mathbf{v}_t \end{bmatrix} \tag{17}$$

[12] Without loss of generality country 1 is the reference country.

[13] The subscript t is omitted to keep the notation simple. In practice the list of non-participating countries varies from one benchmark to another. For non-benchmark years S_{np}remains the same for all t

$$\mathrm{E}(\boldsymbol{\zeta}_t\boldsymbol{\zeta}_t') \equiv \mathbf{H}_t = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \sigma_u^2 \boldsymbol{S}_{np}\boldsymbol{\Omega}_t\boldsymbol{S}_{np}' \end{bmatrix} \tag{18}$$

with σ_u^2 a constant of proportionality, and in (16) the countries are ordered so that the reference country is the first row[14]. In non-benchmark years $\mathbf{S}_{np}$ is an $N-1 \times N$ matrix by definition and $\hat{\mathbf{p}}_t$ is an $N \times 1$ vector of regression predictions for all countries[15].

ii) Benchmark years

$$\mathbf{y}_t = \begin{bmatrix} 0 \\ \mathbf{S}_{np}\hat{\mathbf{p}}_t \\ \tilde{\mathbf{p}}_t \end{bmatrix}, \ \mathbf{Z}_t = \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_{np} \\ \mathbf{S}_p \end{bmatrix}, \ \mathbf{B}_t = \begin{bmatrix} 0 \\ \mathbf{S}_{np} \\ \mathbf{0} \end{bmatrix}, \ \boldsymbol{\zeta}_t = \begin{bmatrix} 0 \\ \mathbf{S}_{np}\mathbf{v}_t \\ \mathbf{S}_p\xi_t \end{bmatrix} \tag{19}$$

$$E\left(\boldsymbol{\zeta}_t\boldsymbol{\zeta}_t'\right) \equiv \mathbf{H}_t = \begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_u^2 \boldsymbol{S}_{np}\boldsymbol{\Omega}_t\boldsymbol{S}_{np}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_\xi^2 \boldsymbol{S}_p\boldsymbol{V}_t\boldsymbol{S}_p' \end{bmatrix} \tag{20}$$

$\tilde{\mathbf{p}}_t$ is an $N_t \times 1$ vector of benchmark observations; $\hat{\mathbf{p}}_t$ is an $N \times 1$ vector of regression predictions for all countries. Again, σ_{u}^2 and σ_ξ^2 are constants of proportionality and the first row is the reference country.

Equations (12) and (16), together with the matrix definitions (17) to (20), constitute the 'transition' and 'observation' equations, respectively of a state space model for the unobservable 'state vector', $\boldsymbol{\alpha}_t$.

Given the unknown parameters, $\boldsymbol{\theta}$, ϕ, σ_u^2, σ_η^2, σ_ξ^2 and the distribution of the initial vector, $\boldsymbol{\alpha}_0$, under Gaussian assumptions, the Kalman filter computes the conditional (on the information available at time t) mean $\hat{\boldsymbol{\alpha}}_t$, and covariance matrix, $\mathbf{P}_t$, of the distribution of $\boldsymbol{\alpha}_t$. Further, $\hat{\boldsymbol{\alpha}}_t$ is a minimum mean square error predictor (MMSE) of the state vector, $\boldsymbol{\alpha}_t$. Even when Gaussian assumptions do not hold, the Kalman filter is still the optimal predictor in the sense that it minimizes the mean square error within the class of all linear predictors (see Harvey [1989], pp. 100-12, Durbin and Koopman [2001] Sections 4.2 and 4.3).

4 Special Features and Properties of the Method

We provide several analytical results including the important result that the constructed series are invariant to the reference country and in certain special cases of the model they are weighted averages of previous observations (benchmarks and regression predictions).

[14] The inclusion of the reference country constraint is a necessary condition for invariance of the results to the chosen reference country.

[15] For invariance to hold it is necessary that the observation for participating countries in benchmark years be the ICP benchmark observations. The estimation of $\boldsymbol{\theta}$, to produce $\hat{\boldsymbol{p}}_t$, is based on all N countries in the sample. See Appendix A and Section 5.2 for details.

The econometric approach, and its state-space representation, encompass a number of models sought by practitioners and international organizations. In this section we demonstrate the versatility of the model by showing how we can adapt the model to suit different scenarios.

4.1 Constraining the model to track benchmark PPPs

As the ICP is the main source of $PPPs$ for different benchmarks and the respective $PPPs$ are determined using price data collected from extensive price surveys, one may consider it necessary that the econometric method proposed should generate predicted $PPPs$ that are identical to benchmark $PPPs$ in the benchmark years. This can be achieved simply by setting $\sigma_\xi^2 = 0$ in (20). The last line in (19) then becomes a constraint, guaranteeing that predicted $PPPs$ are identical to the corresponding benchmark observations. This particular property of Kalman filter predictions follows from the results presented in Doran, 1992.

4.2 Constraining the model to preserve movements in the implicit GDP deflator

In the currently available series, including the Maddison and PWT, growth rates in real GDP as well as the implicit price deflators preserve the national level movements in prices and real income. As the GDP deflator data are provided by the countries and such deflators are compiled using extensive country-specific data, it is often considered more important that the predicted $PPPs$ preserve the observed growth rates implicit in the GDP deflator[16]. This essential feature can be guaranteed in our work by setting $\sigma_\eta^2 = 0$ in (12). In our work we choose to impose this restriction through the backward filter (smoother, see Section 5.2). It is trivial to show that national level movements in prices are preserved using the formulae for the fixed interval Kalman Smoother (see the Appendix B). Growth rate constrained predictions are demonstrated in the empirical section.

4.3 Flexibility in the use of regression predictions

An important feature of our model is that the predictions generated by our national price levels regression model (and information provided by relevant socio-economic variables through model (4)) can be utilized in both benchmark and non-benchmark periods. This is a more

[16]Preserving movements in the implicit deflator will ensure that the growth rates in GDP at constant prices (real) and growth in real per capita income reported and used at the country level are preserved in the international comparisons.

general framework than the approach presented in RRD where the regression predictions are used only for non-participating countries in benchmark years. *This is where this paper extends and presents a general form to the approach suggested in* RRD. In the case of RRD the *algorithm updates predictions between benchmarks using only movements in deflators*. Results obtained under this simplified model are also presented in Section 6.

4.4 Kalman Filter predictions as 'weighted averages' of benchmark year only predictions

In this sub-section we address the question of whether there are conditions under which our predictions can be interpreted as weighted averages of extrapolations from different benchmarks. We identify a set of sufficient conditions for this result to hold.

If there are $M+1$ benchmark years ($j = 0, \ldots, M$)[17], applying growth rates to benchmark $PPPs$ (along with extrapolation to non-participating countries) will produce $M+1$ different panels of PPP estimates. Faced with the dilemma of which panel to use, two possible approaches (of many) would be to: (a) use the panel based on the most recent benchmark year[18]; or (b) to take a weighted average of the $M+1$ different panels. PWT6.3 as well as the Maddison series use the approach in (a). Under (b), there is also a need to specify the weights given to different benchmarks..

Proposition 1: If $PPPs$ for benchmarks from the ICP and regression extrapolations are used only in the benchmark years and if national deflators are used for updating $PPPs$, then the predicted panel of PPP estimates produced by our approach is a 'weighted average' of the $M+1$ panels discussed above. More specifically, suppose $\overleftrightarrow{\boldsymbol{p}}_{t,j}$ is the vector of predicted PPP in year t obtained by applying growth rates to the jth benchmark. Then, denoting the Kalman Filter predictions under this scenario by $\hat{\boldsymbol{\alpha}}_t$, we have

$$\hat{\boldsymbol{\alpha}}_t = \sum_{j=0}^{M} \Upsilon_j^{(M)} \overleftrightarrow{\boldsymbol{p}}_{t,j} \tag{21}$$

where the $\Upsilon_j^{(M)}$ are positive definite matrices such that $\sum_{j=0}^{M} \Upsilon_j^{(M)} = \mathbf{I}_N$

It is in this sense the prediction in (21) is considered a 'weighted average' although it is not generally true that the elements of $\hat{\boldsymbol{\alpha}}_t$ are a weighted average of those of the $\overleftrightarrow{\boldsymbol{p}}_{t,j}$. In the special case when the measurement errors in growth rates and benchmark $PPPs$ are uncorrelated across countries, we can show that the elements of $\hat{\boldsymbol{\alpha}}_t$ are a weighted average

[17]It is convenient for the algebraic derivations to set the number of benchmarks to $M+1$.

[18]This is the approach used in the WDI publication for 2008 and 2009. All the published figures are anchored on the results from the 2005 benchmark comparison.

of the corresponding elements of the $M+1$ 'benchmark only' panels. Then, we have been able to show that

$$\hat{\boldsymbol{\alpha}}_t = \sum_{j=0}^{M} \upsilon_{ii,j}^{(M)} \overleftrightarrow{\boldsymbol{p}}_{t,j} \tag{22}$$

where, $\upsilon_{ii,j}^{(M)}$ is the ith diagonal element of the matrix $\Upsilon_j^{(M)}$, $\upsilon_{ii,j}^{(M)} > 0$ and $\sum_{j=0}^{M} \upsilon_{ii,j}^{(M)} = 1$.

Here we note that the Kalman filter predictions automatically produce a set of weights, $\upsilon_{ii,j}^{(M)}$, for the averaging process. Elements $\upsilon_{ii,j}^{(M)}$ can be interpreted as reflecting the reliability of the jth benchmark. *A proof of the proposition has been provided in the Appendix of RRD.*

The above result has been explored by Harvey and Koopman, 2000 and Koopman and Harvey, 2003 in the context of univariate time series models. RRD showed that for the state-space in (16) and (7), the Kalman filter predictors are weighted sums of all the corresponding elements of the $M+1$ panels (result in (21)) in general, and weighted averages of the corresponding elements of the $M+1$ 'benchmark only' panels in the special case (result in (22)).

4.5 Invariance of the Predictions PPPs to the Choice of Reference country

In the exposition of our model, we used country 1 as the reference or base country with $p_{1,t} = 0$ for all t. This condition is then incorporated into our model and its state-space representation. For this method to be meaningful, it is necessary that it the results are invariant to the choice of the reference country. A significant contribution of this paper is to provide a proof that our econometric approach is invariant to the choice of the reference country.

Proposition 2: If we denote by $\hat{\boldsymbol{\alpha}}_t^{(1)}$ the Kalman Filter predictions when the reference country is 1 (e.g. the USA), and by $\hat{\boldsymbol{\alpha}}_t^{(2)}$ the Kalman Filter predictions of the state vector when the reference country selected is 2 (e.g. the UK); then

$$\hat{\alpha}_{it}^{(2)} = \hat{\alpha}_{it}^{(1)} - \hat{\alpha}_{2t}^{(1)} \tag{23}$$

The proof is fairly complex and it is presented in the Appendix A. 23shows that the predicted PPP for country i with country 2 as the reference country is equal to the rebased PPP for country i computed using country 1 as the numeraire.

5 Estimation

In order for the Kalman filter and smoothing algorithms to deliver a predictor of the state vector and its covariance matrix, we require estimates of the unknown parameters and a distribution for the initial state vector. The parameters of the state-space system can be estimated using likelihood based methods (Harvey, 1989, pp. 125-46) or Bayesian methods (see for instance Durbin and Koopman, 2001, Koop and van Dijk, 2000, and Harvey et al., 2007). The results presented in this paper are obtained using likelihood based methods (details are provided in Section 5.2). The distribution of the initial state vector, $\boldsymbol{\alpha}_o$, is derived as follows.

5.1 Distribution of the Initial State Vector

We specify a distribution with a non-diffuse covariance for the initial state vector, $\boldsymbol{\alpha}_o$, by making use of equation (5). Let $\mathbf{X}_o$ denote a matrix of data on socioeconomic variables in period $t = 0$, the selected benchmark year. Then we can define,

$$\hat{\boldsymbol{\alpha}}_o = \hat{\beta}_{00} + \mathbf{X}_o\hat{\boldsymbol{\beta}}_s + ln(\mathbf{ER}) \tag{24}$$

where,

$\hat{\beta}_{oo}$ is the intercept; $\hat{\boldsymbol{\beta}}_s$ is the estimated slope coefficient vector which is independent of t; $\hat{\boldsymbol{\alpha}}_o = \begin{bmatrix} \hat{\boldsymbol{\alpha}}_0^{(1)} \\ \hat{\boldsymbol{\alpha}}_o^{(2)} \end{bmatrix}$; $\mathbf{X}_o = \begin{bmatrix} \mathbf{X}_o^{(1)} \\ \mathbf{X}_o^{(2)} \end{bmatrix}$; $\mathbf{X}_o^{(1)}$ and $\boldsymbol{\alpha}_o^{(1)}$ represent the partition containing the observations from participating countries.

Then covariance associated to the prediction of $\boldsymbol{\alpha}_o$ is given by

$$\text{cov}(\hat{\boldsymbol{\alpha}}_o) = \mathbf{P}_o = \hat{\sigma}^2\mathbf{X}_o(\mathbf{X}_o^{(1)\prime}\mathbf{X}_o^{(1)})^{-1}\mathbf{X}_o' \tag{25}$$

We use the expression in (25) to obtain an estimate of the covariance of the initial state vector in the empirical section using data for the year 1985 (ICP Phase V year). The intercept and slope estimates in (24) are obtained using a regression containing the participating countries in 1985.

5.2 Estimation of unknown parameters and completion of the panel

There are two types of parameters to be estimated in the state-space, namely, hyperparameters (associated with the covariance structure, which in our case are: ϕ, σ_u^2 , σ_η^2 , σ_ξ^2), and coefficients associated with explanatory variables. Hyperparameters are estimated by numerical maximization of the likelihood function (in a likelihood based estimation). The parameters $\boldsymbol{\theta}$ (those in the regression partition in (16)) can be included as part of the state vector or estimated by a generalized least squares procedure (GLS) in conjunction with the numerical maximization of the likelihood function. We implement the later, which we denote by KF/GLS as it involves running the Kalman filter through both $\mathbf{y}_t$ and the columns of $\mathbf{X}_t$ (see Harvey, 1989, pp. 130-3 or Section 6.2 of Durbin and Koopman, 2001 for more details of this procedure)[19].

One non-standard feature of the state space form in this paper is that some elements of the observation vector, $\mathbf{y}_t$, the $\hat{\mathbf{p}}_t$, are predictions from the price level regression which has errors that are spatially correlated. The observation vector contains these predictions for non-participating countries in benchmark years, and for countries $i = 2, 3, ..., N$ in non-benchmark years. An initial set of predictions to start our algorithm (described below) is obtained by estimating the price level regression for the unbalanced panel composed of participating countries in all benchmark years. These predictions are updated as part of our algorithm as estimates of the parameter ϕ and vector $\boldsymbol{\theta}$ are obtained using the available information on $\mathbf{X}_t$ and $\mathbf{W}_t$ for *all countries and all time periods*. This is achieved by using the Kalman filter algorithm to combine all the information for the purpose of estimation of these parameters.

We first use the state-space form and Kalman filter to obtain the estimates of the unknown parameters and then the filtering-smoothing step to obtain the complete panel, which is run once estimates of the parameters have been obtained.

Parameter Estimation Algorithm:

Step *1*: Obtain an initial estimate of $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}^0$, by regressing $\boldsymbol{r}_t$ on $\boldsymbol{X}_t$ with the panel of benchmark observations without accounting for spatial errors, see equation (4), and construct an initial prediction, $\hat{\boldsymbol{p}}_{it}^0$, using equation (5). An initial estimate of ϕ can then be obtained by computing the correlation between the OLS residuals and lagged residuals (from the regression in (4)). A choice of starting values for σ_u^2, σ_η^2, σ_ξ^2 is also needed. We use a grid search over the range $1e^{-8}$ to $1e^4$ and check the value of the likelihood by running Step 2 to locate a neighborhood of the global maximum. Denote $\boldsymbol{\gamma} =(\phi, \sigma_u^2, \sigma_\eta^2, \sigma_\xi^2)$ and $\hat{\boldsymbol{\gamma}}^0$ the vector of starting values.

Step *2*: Given a set of values for $\boldsymbol{\gamma}$, an estimate of $\boldsymbol{\theta}$ is obtained by a KF/GLS procedure

[19] The code for the empirical estimations was written by the authors in GAUSS and includes a procedure to evaluate the likelihood function when some of the parameters are obtained by the KF/GLS approach .

(explained below) and conditioning on these estimates, denoted by $\boldsymbol{\theta}^j$, a Newton-Raphson iterative procedure is used to maximize the likelihood function and obtain a new estimate of $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\gamma}}^j$. Thus, a set of MLE estimates of ϕ, σ_u^2 , σ_η^2, σ_ξ^2, $\boldsymbol{\theta}$ are obtained at each iteration j. We note that after the first iteration of this step, the estimates of all the parameters in (4) are based on data for the N countries in the sample and account for the spatial correlation structure of the error through the KF/GLS estimation of $\boldsymbol{\theta}$.

KF/GLS estimation: each of the columns of $\mathbf{X}_t$ and $\mathbf{y}_t$ ($K+1$ vectors of dimension N in each time period) are run through the Kalman filter equations to obtain a set of "innovations", denote by $\mathbf{y}_t^*$, $\mathbf{X}_t^*$. The vector $\mathbf{y}_t$ includes observations $p_{1t}=0$, $\tilde{\mathbf{p}}_t$ and $\hat{\mathbf{p}}_t$ if t is a benchmark year, or $p_{1t}=0$ and $\hat{\mathbf{p}}_t$ if t is a non-benchmark year. The matrix $\mathbf{X}_t$ is $N\times K$. The estimation uses data for all countries which is necessary for the invariance result to hold (see Appendix A). The GLS estimator of $\boldsymbol{\theta}$ is computed by regressing $\mathbf{y}_t^*$ on $\mathbf{X}_t^*$ (we refer to it as KF/GLS and the interested reader may consult Harvey, 1989, pp. 130-3 or Durbin and Koopman, 2001, pp. 122-123 for further details).

Step *3*: Using updated estimates, $\hat{\boldsymbol{\theta}}^j$, we obtain revised estimates of $\hat{\beta}_{0t}^j=\hat{\beta}_{0t}^{j-1}-\hat{\theta}_0^j$ and $\hat{\boldsymbol{\beta}}_s^j=\hat{\boldsymbol{\beta}}_s^{j-1}-\hat{\boldsymbol{\theta}}_s^j$, which are used to obtain an updated $\hat{p}_{it}^j=\hat{\beta}_{0t}^j+\mathbf{x}_{it}'\hat{\boldsymbol{\beta}}_s^j+\ln ER_{it}+\hat{u}_{it}^*$, where $\hat{\mathbf{u}}_t^*=\hat{\phi}\mathbf{W}_t\hat{\mathbf{u}}_t^j$. For invariance to hold the predictions require an adjustment by subtracting the base country's prediction, $\hat{p}_{it}^{adjusted}=\hat{p}_{it}^j-\hat{p}_{1t}^j$ (see Appendix A, Section A.2 for details). With the updated $\hat{p}_{it}^{adjusted}$, we update the relevant elements of $\mathbf{y}_t$.

Step *4*: Repeat 2 and 3 until the change in the estimates of $\hat{\theta}_0^j$ and $\hat{\boldsymbol{\theta}}_s^j$ from $j-1$ to j are sufficiently close to zero. In our empirical implementation this occurs after three to four iterations[20].

Upon convergence of the algorithm, the parameters in equations (16), (12) and associated covariances are replaced by their estimates. The observation equation is now given by:

$$\mathbf{y}_t=\mathbf{Z}_t\boldsymbol{\alpha}_t+\boldsymbol{\zeta}_t \tag{26}$$

with variables defined as follows:

i) Non-benchmark years:

$$\mathbf{y}_t=\begin{bmatrix}0\\ \mathbf{S}_{np}\hat{\mathbf{p}}_t^f\end{bmatrix},\ \mathbf{Z}_t=\begin{bmatrix}\mathbf{S}_1\\ \mathbf{S}_{np}\end{bmatrix},\ \boldsymbol{\zeta}_t=\begin{bmatrix}0\\ \mathbf{S}_{np}\mathbf{v}_t\end{bmatrix} \tag{27}$$

[20]This is expected as the iterations over values of $\boldsymbol{\theta}$ simply update the ordinary least squares estimate of $\boldsymbol{\beta}$ to a maximum likelihood estimate through iterations of the GLS estimator.

$$\hat{\mathbf{H}}_t = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \tilde{\sigma}_u^2 \boldsymbol{S}_{np} \tilde{\boldsymbol{\Omega}}_t \boldsymbol{S}'_{np} \end{bmatrix} \tag{28}$$

where $\hat{\mathbf{p}}_t^f = (\tilde{\beta}_{0t} + \boldsymbol{X}\tilde{\boldsymbol{\beta}}_s + \ln(ER_{it}) + \hat{\mathbf{u}}_t^f) - \hat{p}_{1t}$ is the prediction obtained using the estimates from the algorithm described in Steps 1-4 above; $\hat{\mathbf{u}}_t^f = \tilde{\phi}\mathbf{W}_t\hat{\mathbf{u}}_t^*$; $\tilde{\beta}_{0t}$, $\tilde{\boldsymbol{\beta}}_s$, $\tilde{\phi}$, and $\tilde{\sigma}_u^2$ are the estimates obtained from the algorithm in Steps 1-4.

ii) Benchmark years

$$\mathbf{y}_t = \begin{bmatrix} 0 \\ \mathbf{S}_{np}\hat{\mathbf{p}}_t^f \\ \tilde{\mathbf{p}}_t \end{bmatrix}, \ \mathbf{Z}_t = \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_{np} \\ \mathbf{S}_p \end{bmatrix}, \ \boldsymbol{\zeta}_t = \begin{bmatrix} 0 \\ \mathbf{S}_{np}\mathbf{v}_t \\ \mathbf{S}_p\xi_t \end{bmatrix} \tag{29}$$

$$\hat{\mathbf{H}}_t = \begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\sigma}_u^2 \boldsymbol{S}_{np} \tilde{\boldsymbol{\Omega}}_t \boldsymbol{S}'_{np} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{\sigma}_\xi^2 \boldsymbol{S}_p \boldsymbol{V}_t \boldsymbol{S}'_p \end{bmatrix} \tag{30}$$

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \boldsymbol{\eta}_t \tag{31}$$

where $\tilde{\sigma}_\xi^2$ is the estimate obtained from the algorithm in Steps 1-4, and $\hat{\mathbf{Q}}_t = \tilde{\sigma}_\eta^2 \mathbf{V}_t$.

***Smoothing Algorithm*:**

This step estimates the unknown state vector using the state-space in (26) and (31) where all system parameters have now been replaced by estimates. The Kalman Filter (forward filter) and Kalman smoother (backward filter) algorithms [21] are used to obtain the model's predictions of the smoothed state vector α_{it} (for all i and t), $\mathbf{a}_t$, and its covariance matrix, $\hat{P}_t$.

The completed panel provides predictions of $\alpha_{it} = \ln(PPP_{it})$. To obtain predictions of PPP_{it} we reverse the natural log transformation which provides a median unbiased estimate:

$$P\hat{P}P_{it} = e^{a_{it}} \tag{32}$$

where a_{it} is the corresponding Kalman smoothed element.

The standard errors for the predicted $PPPs$ are computed as follows[22]:

$$se(P\hat{P}P_{it}) = \sqrt{e^{2a_{it}} e^{P_{ii,t}} (e^{P_{ii,t}} - 1)} \tag{33}$$

[21] The interested reader will find the equations of the Kalman filter algorithm in Section A.6 of the Appendix and the equations of the Kalman smoother used in this paper in Appendix B.

[22] The standard errors are computed under the assumption of log-normality of the predictions.

where $P_{ii,t}$ is the ith diagonal element of the estimated smoothed covariance of the state vector, P_t.

6 Empirical Results

In this section we present predicted PPPs derived using the methodology proposed in the paper. The basic data used in the empirical analysis are described in 6.1. We cover the period 1970 to 2005 which includes the 2005 benchmark year. Inclusion of 2005 ICP data allows us to produce predicted PPPs under two scenarios, i.e. first using data that includes 2005 and then replicating the analysis without the inclusion of the 2005 data. We are then able to assess the performance of our model in predicting the 2005 ICP PPPs. Further, we are able to compare our estimates with those reported in the recently released PWT 6.3 which are based on benchmark data up to the 2002 OECD/EUROSTAT comparison but provide predictions for 2005. The aim is to demonstrate the flexibility of the method and provide a comparison of results from our method to the results in PWT 6.3.

6.1 Data compilation and data construction

We use data covering 141 countries over the years 1970 to 2005 from RRD. Detailed description of the data used is available in RRD. Appendix Table A.1 lists the 141 countries included in the study; the currency of each country; and the years each country has participated in the ICP Benchmark comparisons. Out of the 141 countries included, 110 are amongst the 146 countries that participated in the 2005 ICP round. That is, there are 31 countries in our data set that did not participate in the 2005 ICP. Appendix Table A.2 gives definitions and sources of the variables used, while Table A.3 provides some basic descriptive statistics of the variables. The dimensions of the data set were largely determined by data availability. That is, a number of countries were excluded because of missing data (see the notes for Table A.1).

6.1.1 PPP Data

The state variable in the state space model is $ln(PPP_{it})$, and observed values of the state vector are obtained through the PPPs from all the benchmarks conducted since 1975[23]. Thus PPP data are drawn from the early benchmarks of 1975, 1980 and 1985 as well as from more recent benchmark information for the years 1990, 1993, 1996, 1999, 2002 and 2005. Several

[23]We have not included the initial benchmarks of 1970 and 1973 as their coverage was too low and they represent the initial phases of the ICP

features of the *PPP* data are noteworthy. The 1975 benchmark covered 34 countries. The 1980, 1985 and the recent 2005, benchmarks represent truly global comparisons with *PPPs* computed using data for all the participating countries. For the years beginning from 1990 to 2002, data are essentially from the OECD and EU comparisons with the exception of 1996[24]. The 1996 benchmark year again was a global comparison with *PPPs* for countries from all the regions of the world. However, the 1996 benchmark may be considered weaker and less reliable than the 1980, 1985 and 2005 benchmark comparisons as no systematic linking of regional *PPPs* was undertaken[25]. Another related point of interest is the fact that *PPPs* for all the benchmarks prior to 1990 were based on the Geary-Khamis method and *PPPs* for the more recent years are all based on the EKS method of aggregation.[26]In the current empirical analysis, we have not made any adjustments to the *PPP* data but making the series comparable through the use of the same aggregation methodology is part of our ongoing research program.

6.1.2 Socio-Economic Variables included in the Regression

The variables used in in the regression model for national prices come under two categories[27]. We use a set of dummy variables designed to capture country-specific episodes that may influence the exchange rates or *PPPs* or both as well as fixed effects. The second set of variables are structural, which are identified from the works of Kravis and Lipsey, 1983, Heston et al., 2006, Clague, 1988, Bergstrand, 1991, 1996 and Ahmad, 1996.

6.1.3 *Variance Specification*

Accuracy of benchmarks and national accounts' growth rates

The heteroskedastic disturbance in equations (2) and (7) are designed to account for the accuracy and reliability of PPPs and the national growth rates through (11). We assume that their reliability is directly related to the real GDP per capita. We make use of exchange rate converted per capita incomes to overcome the problem of possible endogeneity arising out of the use of PPP converted exchange rates. These data are drawn from UN sources.

[24]We are indebted to Ms Francette Koechlin (OECD) for providing ICP benchmark data for these years. *PPPs* for those countries which joined in the Euro zone, the pre-Euro domestic currencies were converted using the 1999 Irrevocable Conversion Rates (Source: `http://www.ecb.int/press/date/1998/html/pr981231_2.en.html`). The irrevocable conversion rate of the drachma vis a vis the euro was set at GRD 340.750. Source: `http://www.bankofgreece.gr/en/euro`.

[25]The weaker quality of the 1996 benchmark is the main reason why the 1996 results were never published formally by the World Bank.

[26]This was brought to our attention by Steve Dowrick of the Australian National University.

[27]See RRD for further details.

Given the systematic nature of deviation of exchange rates from PPPs, use of exchange rate converted per capita GDP is likely to magnify differences in per capita incomes.

Measuring spatial correlation*:*

An important contribution of our study which is a major departure from all the previous empirical studies using a national price levels regression, is the assumption that the disturbances are spatially autocorrelated. Typically, spatial autocorrelations are modelled with weight matrices that reflect either geographical contiguity or the strength of trade relations. In this study we use a comprehensive measure of spatial autocorrelation. Details of our approach are presented next.

A range of variables could be used in modelling spatial autocorrelation and in designing the spatial weights matrix. ** et al., 2010 discuss the specification of three alternative weight matrices and examine the sensitivity of the results. Based on their analysis, a measure of socio-economic distance is constructed by extracting a common factor (through principal components analysis). For each country, the model combines trade closeness, geographical proximity, and cultural closeness. We briefly summarize the approach here.

Variables included in the measure of socio-economic distance

(i) *Trade closeness* (TC) is measured as the percentage of bilateral trade between each country and all others in the sample (compiled using data from Rose, 2004 and IMF Trade Directions).

(ii) *Geographical proximity* is measured by a series of dummies for *border* (both land and sea proximity) (B), and *regional membership* (such us Asia pacific region, Europe, south America, north and central America, sub Saharan Africa, middle east) (RM). The data were constructed using Atlas, CIA factbook and individual country references.

(iii) *Cultural and colonial closeness* dummies are used for *common language* (CL) and *common colonial history* (CH). The data were constructed from the CIA factbook and individual country references.

Construction of the distance score

The steps involved in the estimation of the "distance" between pairs of countries are as follows:

1) A principal components (PC) model was estimated for each country. That is, for any country i we can define an $((N-1)\times 5)$ matrix of observed pairwise measures (ij for $j \neq i = 1, ..., N$) on the five variables (TC, B, RM, CL and CH). The estimation was conducted for the years 1970, 1975, 1980, 1985, 1990, 1995, 2000 and 2005 to account for the changing patterns in bilateral trade over time. That is, for each time period we estimate 141 (N) principal components models.

2) The first PC explaining the maximum variability is identified as the *common factor.*

3) A factor score is computed using the estimated factor loadings and the data. These scores are not bounded; therefore, they are rescaled to prepare the *proximity matrix* with elements $S_{ij} = \left[\frac{f_{ij}-f_{i,\min}}{f_{i,\max}-f_{i,\min}}\right]$;

where, $f_{i,\min}$, $f_{i,\max}$ and f_{ij} are respectively the minimum factor score value, the maximum factor score value and factor score of country i in relation to j. These rescaled factor scores are in the range of 0 to 1, and if $i = j$ the rescaled value is zero.

The distance or proximity score is assumed to be constant within the five yearly intervals but vary over time from 1970 to 1974, 1975 to 1979, and so on.

Construction of the Weights Matrix

The proximity matrix is transformed into a row stochastic matrix $\boldsymbol{W}_t$ (i.e. rows add up to one) by simply dividing each proximity score within a row (which represents a country) by the sum of that row.

6.2 Results

We divide our presentation into two parts. First, we present and discuss the testing for spatial correlation (cross-sectional dependence) and then discuss the estimates of the parameters as well as the goodness of fit with special reference to those in the price level model. Second, using a selection of five countries we present and discuss the *PPPs* predictions obtained using our method and highlight its flexibility as well as compare its out-of-sample prediction ability and to the predictions of PWT6.3.

6.2.1 Testing for Spatial Correlation and Estimates of Parameters

Testing for Spatial/Cross-Sectional Dependence

The price level model of equation (4) is an unbalanced panel with fixed time effects. The available data to test the residuals of this model correspond to those years when there has been either an ICP or OECD/EUROSTAT benchmark comparison (that is, 1975, 1980, 1985, 1990, 1993, 1996, 1999, 2002 and 2005 in our sample). Commonly used tests for spatial dependence (LM, Moran's Statistic) require the specification of a spatial model (that is $H_o : \phi = 0$ in eq. (8)), and therefore are dependent on the specification of the spatial weights matrix, $\boldsymbol{W}_t$. An alternative strategy is to use a robust test for cross-sectional dependence, such as the CD Test by Pesaran, 2004, which does not require the specification of a spatial model. The test is based on simple averages of all pair-wise correlation coefficients of the OLS residuals from the individual regressions in the panel. For the case of unbalanced panels the CD test takes the following form (the reader is referred to Section 9 of Pesaran, 2004 for more details):

$$\hat{\rho}_{ij} = \frac{\sum_{m \in T_i \cap T_j} (\tilde{u}_{im} - \bar{u}_i)(\tilde{u}_{jm} - \bar{u}_j)}{\left[\sum_{m \in T_i \cap T_j} (\tilde{u}_{im} - \bar{u}_i)^2\right]^{1/2} \left[\sum_{m \in T_i \cap T_j} (\tilde{u}_{jm} - \bar{u}_j)^2\right]^{1/2}} \tag{34}$$

where $\hat{\rho}_{ij}$ is the correlation coefficient between country i and j; T_i is the set of benchmark years where country i has participated in the ICP; $\tilde{u}_{im}$OLS residual for country i in benchmark year m; $\bar{u}_i = \dfrac{\sum_{m \in T_i \cap T_j} \tilde{u}_{im}}{\mathrm{T}_{ij}}$; T_{ij} is the number of elements in $T_i \cap T_j$.

The CD statistic for the unbalanced panel is given by $CD = \sqrt{\dfrac{2}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sqrt{\mathrm{T}_{ij}} \hat{\rho}_{ij} \right)$

Under the null hypothesis of no cross sectional dependence, $CD \sim N(0, 1)$. In our case, the computed value of the CD test is -100.9, and therefore the null hypothesis is rejected at all levels of significance[28].

Parameter Estimates

Table 1 presents a series of estimated models that are then used to construct the *PPPs* series for all 141 countries. Estimates for different models are presented, but due to space constraints the time intercept dummy estimates are not included. Panel 1 is the national price level regression model (equation (4)) estimated by least squares assuming non-spatial errors and data covering all the benchmarks since 1975 to 2005. This model is used to produce the initial predictions to start the estimation algorithm (refer to Section 5.2).

Panel 2 are the estimates from the state-space model produced without restrictions on any of the parameters in the model. The price level regression model is assumed to have spatial errors. The estimate of the spatial parameter is 0.59 and it is statistically significant. The covariance proportionality parameters associated with the error in the growth rates, regression predictions and ICP benchmarks are estimated to be 6.6, 4.5, and 0.8, respectively, and they are all statistically significant. As the *PPP* series obtained from the Kalman filter and smoother are not constrained to track ICP benchmarks or growth rates, *PPP* predictions from this model are labeled with the postfix "UN."

Panel 3 is the state-space model estimates obtained by restricting the parameter that controls the error in ICP benchmarks to zero, i.e. $\sigma_\eta^2 = 0$. The spatial parameter as well as the parameters associated with errors in the growth rates and regression predictions are statistically significant. The log-likelihood of this model is lower than that of the model in

[28] LM tests for spatial correlation were computed for three alternative specifications of the weight matrix and results can be found in ** et al., 2010

Panel 2. These can be compared by a likelihood ratio test as the model in Panel 3 is a restriction of the model in Panel 2. The computed LR value is significantly different from zero and therefore the restriction that benchmarks do not suffer from measurement error is rejected. As the Kalman smoothed series produced by this model are constrained to track ICP benchmarks, the *PPP* predictions from this model will be labeled with the postfix "CON."

Panel 4 is a simpler form of the general model in Panel 2 in that in non-benchmark years no regression prediction information is used (this is the model used in RRD). The regression predictions are used only in benchmark years for non-participating countries (see equations (29) and (30)). For the years between benchmarks the only information included is the temporal movement in prices through the transition equation (31). It has been shown that the model's estimates are weighted averages of the observed ICP benchmarks for countries that participated in all benchmarks, a weighted average of the combination of the ICP benchmarks and regression predictions for countries that only participated in some of the ICP benchmarks, or a weighted average of the regression predictions from the national price level model for those countries that never participated in an ICP benchmark (see Section 4.4). The weights decrease inversely with time so that older observations are weighted less (see Appendix in RRD). Both benchmarks and growth rates are assumed to be measured with error as in Panel 2. The value of the likelihood functions is higher than that of Panel 2 although the two models are not strictly nested. The smoothed PPPs series produced from this model will be labeled with the postfix "No Reg." However, the smoothed predictions (presented in the next section) have standard errors that are larger than those produced from the model in Panel 2 in most cases.

Panel 5 refers to the model which allows a comparison of our predictions to those available from PWT 6.3. As the latter were produced without using data from the 2005 round of the ICP, we estimate our model for the time period 1971-2005 as before; however, the year 2005 is treated as a non-benchmark year. Identical to the case of Panel 2, all sources of measurement errors are allowed as parameters are not restricted. An equivalent regression to that in Panel 1 is run to obtain starting values. We will label these *PPP* predictions by "No05."

[Table 2]

6.2.2 PPP Predictions

In this section we present the predictions obtained using the models presented in the previous section. The performance of models without spatial errors was very poor and therefore those results are not presented or discussed in this paper. A comparison of the in-sample and out-of-sample prediction performance of the model with and without spatial errors can be found in ** et al., 2010, Section 5.2.2.

Two sets of predicted *PPP* series can be computed for each model depending on how the Kalman filtered predictions obtained from the above models are smoothed. The first set is obtained by smoothing the Kalman filtered predictions without imposing any extra restrictions. A second set of predictions is obtained under a form of the smoother that insures the resulting series tracks the price deflator movements published for each country (see Appendix B for details). The latter will be distinguished from the first by the postfix "GRC." Thus, GRC series are those when the predictions track the movements in the implicit deflators.

The series labeled "CON" are those obtained from the model in Panel 3 and they are constrained to track the observed benchmarks. The corresponding standard errors for participating countries in benchmark years are zero. However, standard errors for non-benchmark years are larger than those estimated by the unrestricted version of the model (Panel 2).

Tables 3-7 and Figures 1-3 summarize our results. As the complete panel of PPPs covering 35 years and 141 countries is very large , we have chosen five countries for illustration, viz., Australia, China, India, Nigeria and Honduras[29]. Australia is an OECD country and has participated in all benchmark comparisons since 1985. Results for Australia are representative of results for developed countries that have consistently participated in most of the global as well as OECD comparisons; Australia represents a case when all sources of available information (national accounts and benchmark data) seem to provide a consistent picture. In contrast, China participated in ICP benchmark comparisons for the first time in 2005. India had participated in earlier benchmarks; however it had not participated since 1985 and has again participated in the 2005 round. Nigeria had participated in the global comparisons of 1975, 1980, 1985, 1996 and the 2005 round. Honduras had participated in the 1980 comparison and it is one of the countries in the sample that did not participate in the 2005 round.

Australia

Table 3 and Figure 1 present the *PPP* predictions for Australia. We note the consistency between the series where the movements in the implied price deflator are maintained (labeled GRC) and the ICP *PPP* benchmarks, specially since 1990, across all estimated models (see Table 3). In Figure 1 it is clear that from 1985 onwards all alternative *PPP* series are within the two-standard errors band generated by the model in Panel 5 (without 2005 benchmark). The standard error for the 2005 prediction is AUD 0.05 and reduces to AUD 0.01 when the

[29]Detailed results for other countries are available from the authors upon request.

2005 benchmark is actually included. We also note that the "GRC" predictions appear to track the benchmark PPPs automatically without any additional constraints.

[Table 3 and Figure 1]

China

Figures 2 and Table 4 present our predictions for China. A few important points can be made. First, the predictions that have not been smoothed to follow the published GDP Deflator movements (see series labeled PPP-UN) differ substantially from the series obtained when this is imposed (PPP-UN-GRC) specially before 1990, indicating that internationally available data on socio-economic variables for China, especially for the years before 1990, provide a different picture than what is available in the form of movements in the published GDP Deflators. Further, and as expected, the standard error of the estimates generated from Panel 5 (without the 2005 benchmark) is very large, Yuan 1.684. The standard error reduces to Yuan 0.092 when the 2005 data are included (to Yuan 0.103 for the predictions from the model in Panel 4). We compare our China predictions with PWT 6.3. PWT6.3 provides two alternative series for China, China-Version 1 and -Version 2. China -Version 1 uses the official growth rates published by Chinese national accounts, while the second uses Maddison and Wu, 2007 (see the Papers section of the PWT site). We show both in Figure 2 and version 1 in Table 4. Our PPP predictions are comparable to PWT6.3 version 1. The performance of our method in predicting the 2005 *PPP* value is substantially better than that of PWT 6.3 as our model was predicting the *PPP* for China in 2005 to be Yuan 3.09, while PWT 6.3's prediction was Yuan 2.23. The 2005 ICP benchmark was Yuan 3.45 . The main reason for the superior performance of our model predictions is that we combine regression predictions with other sources. We believe that our panel regression model has considerable power to extrapolate PPPs.

[Table 4 and Figure 2]

India

The Indian case is different from that of China (refer to Table 5 and Figure 3). India participated in several benchmarks; however, its last participation before 2005 was in 1985. The differences between PPP-UN and PPP-UN-GRC are large, as in the case of China, for the earlier part of the sample. The PPP-UN is close to the benchmark observations as expected; however, it is clear that the movements implied by the published GDP deflator are inconsistent with earlier benchmarks (see PPP-UN-GRC). For instance, for 1985, the benchmark was Rupee 4.667, while the estimated value when growth rates implied by the GDP Deflator are maintained is Rupee 5.952. The *PPP* series derived from the model without the 2005 benchmark is closer to the actual observation in the 2005 round (Rupee 14.670) than that of PWT6.3. The PWT6.3 estimate is Rupee 9.540, while our estimate is Rupee 10.376 (standard error of Rupee 5.331). The large standard error arises because the last available ICP benchmark for India is 1985 and there are some inconsistencies between the information from the socioeconomic variables and the GDP deflator, which introduces the uncertainty shown in the standard errors. The inclusion of the 2005 benchmark reduces the standard error to Rupee 0.502 (using the model in Panel 2).

[Table 5 and Figure 3]

Nigeria

Nigeria participated in four benchmarks, 1980, 1985, 1996 and 2005 (results are presented in Table 6). As in the case of India, it is clear that the growth rates implied by the published GDP deflator is inconsistent with earlier benchmarks. The results for Nigeria highlight the inconsistencies between the information contained in socio-economic variables, GDP deflators and earlier benchmarks. The prediction from the model in Panel 2 (UN) is Naira 60.096 (with a standard error of 2.894), which is much lower than the prediction without the 2005 benchmark, Naira 65.97 (with a standard error of Naira 30.9). The ICP benchmark for 2005 is Naira 60.00. The PWT6.3 for 2005 is Naira 55.28 which is much lower than our comparable estimate.

[Table 6]

Honduras

Honduras is one of the countries that did not participate in the 2005 round of the ICP. As no Central American country participated in this benchmark comparison, the available information for the region is only that from socio-economic variables and GDP deflators. Table 7 shows the results. The estimated series from the model without the 2005 benchmark data predicts the 2005 *PPP* to be Lempira 10.596 (standard error of 5.052). The PWT6.3 for 2005 is Lempira 7.612 which is substantially lower. The predicted series from the model in Panel 2 (UN), when the 2005 ICP benchmark values are used by the model, is Lempira 10.337 for 2005 with a standard error of 4.918. It is also worth noting that the benchmark constrained model has the largest standard errors, Lempira 6.410 for 2005.

[Table 7]

6.3 Discussion

The results for a handful of countries were used to illustrate in the previous section; however, from the overall empirical results and constructed panel (available from the author's) we can provide the following summary:

1. For the majority of countries, the *PPP* predictions are improved by the inclusion of regression information both in benchmark and non-benchmark years in that the standard errors are smaller if all the information from regression predictions is used instead of the simplified version in RRD which excludes regression information in non-benchmark years. For a small group of developed countries that have consistently participated in the ICP and OECD/Eurostat benchmark comparisons, the inclusion of the regression information does not improve the predictions, as expected, and it might result in slightly larger standard errors when the regression information is included. However, there are only 23 countries in this group.

2. The use of the full state-space model is justified when comparing the predictions from our model without the inclusion of the 2005 benchmark data to those by PWT 6.3 and the actual benchmark values produced by the ICP for 2005. Predictions from our approach when all sources of information (all benchmarks and regression predictions), except for the 2005 benchmark values are included are much closer to those found through the 2005 ICP round than those by PWT6.3 for most countries (see China and India, Tables 4 and 5). Furthermore, and as expected, the difference between the predictions of our method, the ICP benchmarks and PWT 6.3 for countries such as Australia are minimal especially after the mid-1990s. This means that if efficient predictions from a smoothing method that makes use of all past benchmarks had been applied, it would have resulted in much smaller adjustments to the 2005 ICP results published early in 2008.

3. The strongest contribution of the 2005 ICP round has come in the form of a reduction in uncertainty, which is very clear by comparing the size of the standard errors for the models with and without 2005 benchmark data included. This reduction is in part due to the impressive cover of 146 countries achieved in the 2005 ICP round.

7 Conclusions

The econometric methodology suggested in the paper for the construction of a consistent panel of purchasing power parities represents a major step forward as it advocates a transparent and coherent approach. The approach suggested is designed to make use of all the principal and auxiliary information available for the purpose of extrapolation of the International Comparison Program (ICP) benchmarks. Existing approaches to the construction of panels of PPPs are two-step methods, while the new method is a single step method. The econometric model is expressed as a state-space model as the problem of estimating PPPs is one of signal extraction. The paper demonstrates that the new approach is flexible in that it can be used to consider a number of scenarios including restrictions on some variance parameters to generate extrapolations that track the observed ICP PPPs in benchmark years or the implied price movements over time for individual countries. A proof of the invariance of the resulting predictions to the choice of reference country is provided. Further, this is the first available approach to producing not only a panel of PPPs, but also associated standard errors that can be incorporated into any further modelling using these estimates.

The methodology proposed is applied to a large data set covering 141 countries and a thirty-five year period 1970 to 2005 for generating predictions. The results from the empirical estimation are illustrated through the PPP series generated for a selected group of countries, including China, India, Australia Nigeria and Hounduras, to examine the plausibility of the extrapolations. The results from the new methodology are contrasted with the published PPPs from the Penn World Table's latest Version 6.3. The results are satisfactory and very encouraging. Further analysis and study of the results for all the 141 countries is currently underway and it is expected that the full panel of PPPs will soon be released for public use.

Tables and Figures

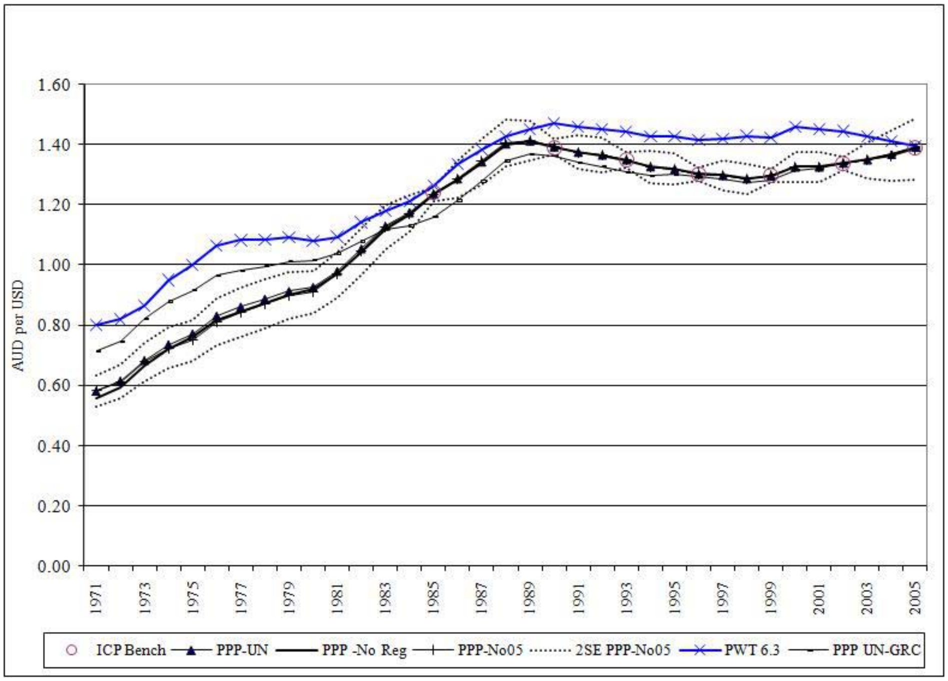

Figure 1: Australia. Extrapolated PPPs using models in Panels 2,4 and 5 and the constrained smoother

ICP Bench: PPPs in benchmark years from the ICP; PWT 6.3: PPPs from Version 6.3 of the Penn World Tables; PPP-UN: predicted PPPs from our model in Panel (2) of Table 2 (includes socio-economic weighted spatial errors and the 2005 ICP benchmark data); PPP -No Reg: predicted PPPs from the model in Panel (4) of Table 2 (includes socio-economic weighted spatial errors and the 2005 ICP benchmark data, but the regression predictions are only used in benchmark years); PPP-No05: predicted PPPs from the model in Panel (5) of Table 2 (includes socio-economic weighted spatial errors but 2005 is treated as a non-benchmark year) and are useful in assessing the robustness of predictions from our model with respect to the inclusion of the 2005 benchmark data; and PPP-UN-GRC: predicted PPPs from the model in Panel (2) (includes socio-economic weighted spatial errors and the 2005 ICP benchmark data but the smoothing is using the restricted version in Appendix B).

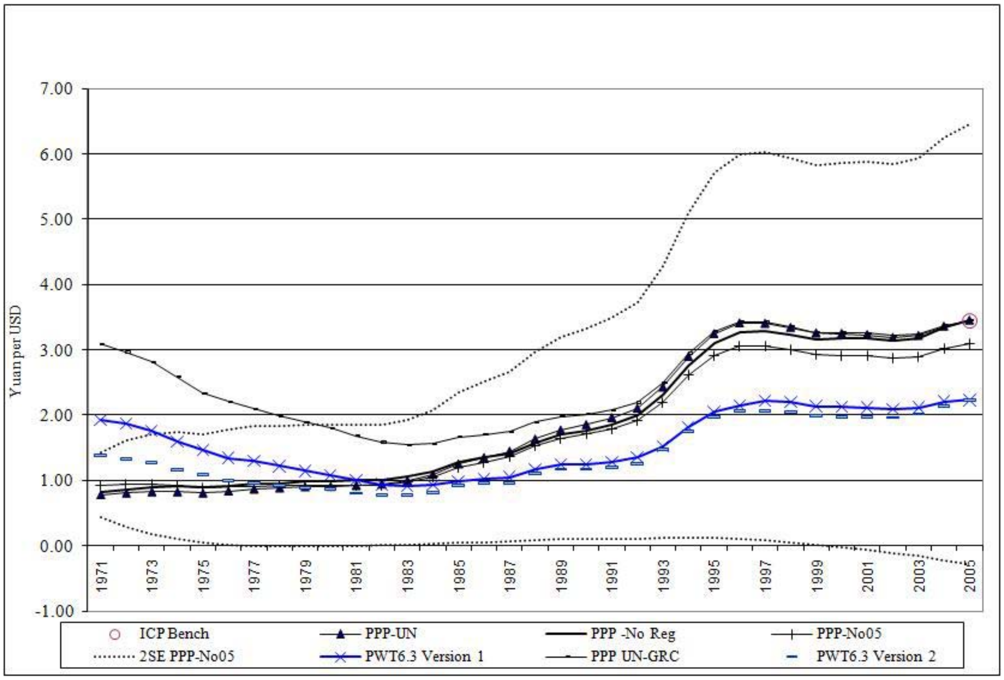

Figure 2: China. Extrapolated PPPs using models in Panels 2,4 and 5 and the constrained smoother.

Note: Legends used here are the same as those used in Figure 1.

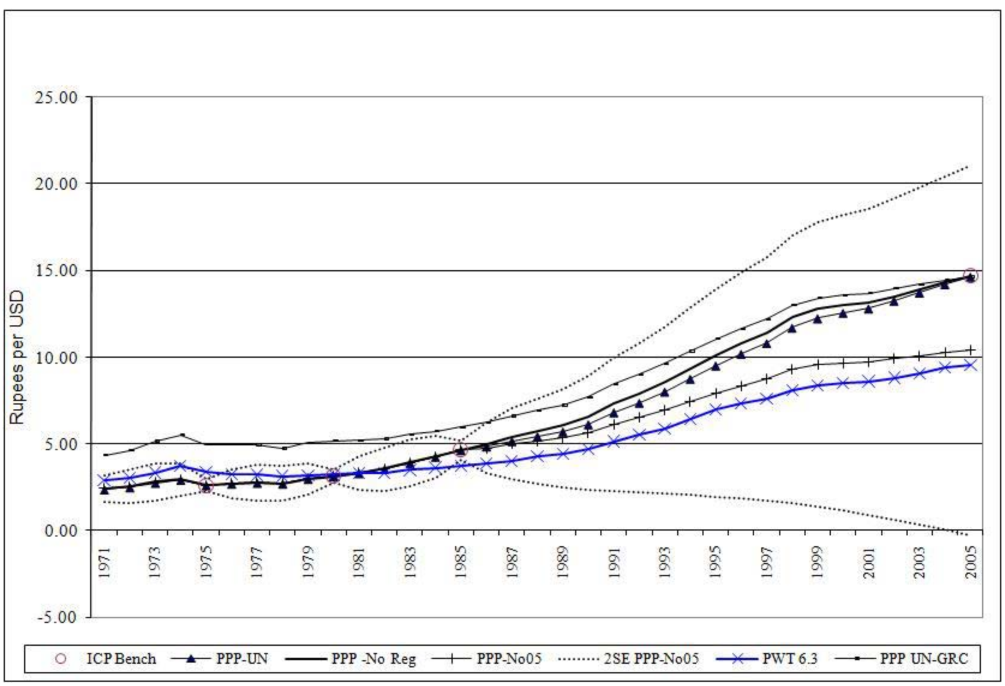

Figure 3: India. Extrapolated PPPs using models in Panels 2,4 and 5 and the constrained smoother.

Note: Legends used here are the same as those used in Figure 1.

Table 1: Number of Countries Participating in ICP (various Phases) and OECD/EuroStat Only Comparisons

ICP Phase	Benchmark Year	Total No. of Participating Countries	Included in This Study
I	1970	10	
II	1973	16	
III	1975	34	33
IV	1980	60	58
V	1985	64	56
OECD/EUROSTAT	1990	24	24
VI	1993/1996	117	112
OECD/EUROSTAT	1999	28	28
OECD/EUROSTAT	2002	28	28
VII	2005	146	110

Source: Asian Development Bank, 2007 and authors

Note: ICP coverage in different phases is global with participating countries from different regions.

Table 2: Parameter Estimates Under Alternative Specifications

	Regression		State Space Model							
	Without Spatial Errors (Panel1)		Benchmark Unconstrained (Panel 2)		Benchmark Constrained (Panel 3)		No Regression In Non-Benchmark Years (Panel 4)		2005 Not a Benchmark Year (Panel 5)	
Variable	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
D_anz	-0.770	0.221	-0.443	0.394	-0.628	0.468	0.365	0.730	-0.491	0.457
D_asean	0.016	0.080	0.075	0.281	0.189	0.356	-2.169	0.481	0.092	0.307
D_cac	-0.029	0.155	0.221	0.278	0.052	0.349	1.376	0.505	0.274	0.307
D_cafrica	0.101	0.116	0.033	0.321	-0.190	0.400	-1.376	0.532	0.348	0.353
D_eafrica	0.118	0.094	0.090	0.283	-0.123	0.354	-0.269	0.466	0.322	0.311
D_euro	0.092	0.045	0.104	0.170	0.005	0.199	0.699	0.333	0.173	0.201
D_mena	0.045	0.073	-0.041	0.194	-0.028	0.240	-1.376	0.361	0.063	0.219
D_mercsr	-0.081	0.082	0.720	0.274	0.757	0.347	-0.476	0.518	0.881	0.304
D_nafta	-0.243	0.086	-0.023	0.305	-0.129	0.369	-0.993	0.611	-0.011	0.352
D_safrica	0.066	0.122	-0.052	0.302	-0.262	0.372	1.019	0.539	-0.002	0.338
D_scucar	0.228	0.148	0.293	0.261	0.328	0.316	0.246	0.467	0.279	0.295
D_spr	0.632	0.206	0.925	0.302	0.987	0.392	2.451	0.536	0.896	0.326
D_usd	0.073	0.069	0.569	0.138	0.608	0.165	0.934	0.253	0.550	0.159
D_wafrica	0.256	0.089	-0.551	0.269	-0.829	0.337	-0.760	0.461	-0.139	0.297
Agedep	0.365	0.174	-0.258	0.571	-0.299	0.661	1.398	0.842	0.272	0.641
Agvagun	-0.009	0.002	-0.019	0.007	-0.018	0.008	-0.001	0.010	-0.014	0.007
Tractorpw	0.094	0.061	0.159	0.245	-0.012	0.280	0.575	0.410	0.071	0.281
Labpop	-0.003	0.003	-0.013	0.011	-0.008	0.013	0.018	0.016	-0.013	0.013
Life	-0.006	0.004	-0.007	0.012	-0.004	0.014	-0.003	0.020	0.009	0.013
Literate	2.1E-04	1.4E-04	-4.0E-04	4.2E-04	-4.9E-04	4.9E-04	0.004	0.001	-3.9E-04	4.6E-04
Ntrvag2	-0.004	0.003	-0.012	0.008	-0.013	0.009	0.018	0.010	-0.012	0.009
Blackind	-0.002	0.003	-0.006	0.006	-0.009	0.007	-0.004	0.011	0.003	0.007
Expg	0.001	1.8E-04	0.003	0.001	0.003	0.001	0.003	0.001	0.002	0.001
Phones	5.0E-06	7.0E-06	-5.5E-05	2.3E-05	-6.3E-05	2.6E-05	-2.6E-04	3.6E-05	-4.4E-05	2.6E-05
Radpccn	-0.004	0.001	-0.004	0.005	-0.003	0.005	-0.020	0.006	-0.005	0.005
Rurpop	3.3E-05	5.4E-05	-9.5E-05	1.7E-04	-5.4E-05	1.9E-04	-0.001	2.8E-04	6.0E-06	1.9E-04
Tradegun	-2.1E-04	0.002	2.4E-05	0.003	0.001	0.004	-0.008	0.006	-0.004	0.004
Manufexp	-2.4E-04	0.001	0.002	0.002	0.001	0.003	0.016	0.004	0.003	0.003
Manufimp	0.003	0.001	0.004	0.004	0.006	0.004	-0.020	0.006	0.005	0.004
$R^2 \mid lnL$	0.737		-1.3E+7		-1.03E+7		-1.46E+7		-1.4E+04	
No Benchmark Samples	449		449		449		449		339	
σ^2_η			7.000	0.005	12.00	0.003	8.000	0.005	7.000	0.420
σ^2_u			4.500	0.104	6.118	0.145	6.000	0.060	6.500	0.045
σ^2_ξ			0.800	0.002	0.000	–	2.000	0.003	0.800	1.0E-16
ϕ			0.700	0.020	0.930	0.008	0.930	0.004	0.550	9.0E-04

Note: Estimates of the time dummy intercepts are not presented due to space constraints.

Table 3: Predicted PPP Series for Australia

YEAR	ER	ICP	UN	SE	UN-GRC(a)	CON	SE	No Reg-GRC(b)	No05-GRC(c)	PWT6.3
1971	0.883		**0.581**	0.026	**0.712**	**0.595**	0.031	**0.713**	**0.710**	0.800
1972	0.839		**0.614**	0.030	**0.745**	**0.630**	0.035	**0.745**	**0.742**	0.818
1973	0.703		**0.683**	0.034	**0.821**	**0.699**	0.040	**0.821**	**0.818**	0.864
1974	0.697		**0.736**	0.036	**0.878**	**0.749**	0.042	**0.878**	**0.874**	0.949
1975	0.764		**0.772**	0.036	**0.915**	**0.778**	0.042	**0.915**	**0.911**	0.999
1976	0.818		**0.830**	0.041	**0.964**	**0.843**	0.048	**0.964**	**0.960**	1.063
1977	0.902		**0.861**	0.043	**0.981**	**0.876**	0.050	**0.981**	**0.977**	1.081
1978	0.874		**0.887**	0.043	**0.995**	**0.901**	0.051	**0.995**	**0.991**	1.082
1979	0.895		**0.912**	0.042	**1.011**	**0.922**	0.048	**1.011**	**1.007**	1.091
1980	0.878		**0.925**	0.037	**1.015**	**0.928**	0.043	**1.016**	**1.011**	1.078
1981	0.870		**0.979**	0.040	**1.037**	**0.986**	0.046	**1.037**	**1.033**	1.093
1982	0.986		**1.053**	0.041	**1.077**	**1.064**	0.047	**1.078**	**1.073**	1.141
1983	1.110		**1.128**	0.038	**1.118**	**1.140**	0.044	**1.118**	**1.113**	1.179
1984	1.140		**1.173**	0.031	**1.129**	**1.183**	0.034	**1.130**	**1.125**	1.211
1985	1.430	1.240	**1.235**	0.011	**1.158**	**1.240**	0.000	**1.158**	**1.153**	1.263
1986	1.500		**1.285**	0.031	**1.213**	**1.292**	0.035	**1.214**	**1.208**	1.334
1987	1.430		**1.341**	0.039	**1.277**	**1.348**	0.044	**1.277**	**1.271**	1.383
1988	1.280		**1.400**	0.040	**1.345**	**1.406**	0.046	**1.345**	**1.339**	1.425
1989	1.260		**1.410**	0.034	**1.367**	**1.413**	0.037	**1.367**	**1.361**	1.449
1990	1.280	1.389	**1.391**	0.012	**1.361**	**1.389**	0.000	**1.362**	**1.356**	1.469
1991	1.280		**1.373**	0.030	**1.339**	**1.372**	0.033	**1.340**	**1.334**	1.457
1992	1.360		**1.363**	0.029	**1.325**	**1.364**	0.033	**1.326**	**1.320**	1.450
1993	1.470	1.350	**1.348**	0.011	**1.306**	**1.350**	0.000	**1.307**	**1.301**	1.442
1994	1.370		**1.325**	0.028	**1.295**	**1.326**	0.031	**1.296**	**1.290**	1.426
1995	1.350		**1.317**	0.027	**1.299**	**1.316**	0.030	**1.299**	**1.293**	1.426
1996	1.280	1.299	**1.300**	0.010	**1.291**	**1.299**	0.000	**1.292**	**1.286**	1.414
1997	1.350		**1.296**	0.026	**1.286**	**1.294**	0.029	**1.287**	**1.281**	1.417
1998	1.590		**1.285**	0.025	**1.273**	**1.284**	0.028	**1.274**	**1.268**	1.427
1999	1.550	1.297	**1.296**	0.010	**1.281**	**1.297**	0.000	**1.282**	**1.276**	1.422
2000	1.720		**1.323**	0.026	**1.314**	**1.323**	0.028	**1.315**	**1.309**	1.457
2001	1.930		**1.324**	0.025	**1.320**	**1.324**	0.028	**1.320**	**1.314**	1.449
2002	1.840	1.337	**1.337**	0.010	**1.337**	**1.337**	0.000	**1.337**	**1.331**	1.444
2003	1.540		**1.348**	0.025	**1.349**	**1.347**	0.028	**1.350**	**1.344**	1.426
2004	1.360		**1.365**	0.025	**1.366**	**1.365**	0.028	**1.367**	**1.361**	1.410
2005	1.309	1.390	**1.390**	0.010	**1.390**	**1.390**	0.000	**1.390**	**1.384**	1.393

Growth Rates Preserved SE: *(a)* 0.010 *(b)* 0.011 *(c)* 0.051.

Table 4: Predicted PPP Series for China

year	ER	ICP	UN	SE	UN-GRC(*a*)	CON	SE	No Reg-GRC(*b*)	No05-GRC(*c*)	PWT6.3 *(d)*
1971	2.460		**0.780**	0.206	**3.086**	**0.768**	0.236	**3.083**	**2.765**	1.927
1972	2.250		**0.805**	0.281	**2.961**	**0.779**	0.318	**2.958**	**2.652**	1.869
1973	1.990		**0.824**	0.333	**2.808**	**0.786**	0.370	**2.805**	**2.515**	1.760
1974	1.960		**0.822**	0.363	**2.582**	**0.774**	0.397	**2.579**	**2.312**	1.591
1975	1.860		**0.806**	0.375	**2.331**	**0.750**	0.403	**2.329**	**2.088**	1.464
1976	1.940		**0.836**	0.402	**2.200**	**0.777**	0.430	**2.198**	**1.971**	1.342
1977	1.860		**0.866**	0.423	**2.091**	**0.800**	0.448	**2.089**	**1.873**	1.300
1978	1.680		**0.881**	0.431	**1.980**	**0.807**	0.451	**1.978**	**1.774**	1.218
1979	1.550		**0.899**	0.437	**1.893**	**0.814**	0.450	**1.891**	**1.696**	1.144
1980	1.500		**0.907**	0.434	**1.801**	**0.809**	0.440	**1.800**	**1.613**	1.073
1981	1.700		**0.920**	0.432	**1.685**	**0.818**	0.435	**1.683**	**1.509**	1.000
1982	1.890		**0.942**	0.431	**1.585**	**0.838**	0.434	**1.583**	**1.420**	0.939
1983	1.980		**0.997**	0.443	**1.541**	**0.888**	0.447	**1.539**	**1.380**	0.914
1984	2.320		**1.088**	0.468	**1.558**	**0.969**	0.472	**1.556**	**1.395**	0.929
1985	2.940		**1.250**	0.521	**1.665**	**1.116**	0.526	**1.663**	**1.491**	0.981
1986	3.450		**1.347**	0.543	**1.703**	**1.211**	0.553	**1.702**	**1.526**	1.022
1987	3.720		**1.445**	0.563	**1.741**	**1.308**	0.578	**1.740**	**1.560**	1.048
1988	3.720		**1.630**	0.613	**1.888**	**1.486**	0.633	**1.886**	**1.691**	1.173
1989	3.770		**1.769**	0.638	**1.979**	**1.619**	0.664	**1.977**	**1.773**	1.241
1990	4.780		**1.852**	0.638	**2.013**	**1.702**	0.667	**2.011**	**1.803**	1.246
1991	5.320		**1.958**	0.643	**2.076**	**1.809**	0.676	**2.074**	**1.860**	1.280
1992	5.510		**2.104**	0.657	**2.190**	**1.954**	0.695	**2.188**	**1.962**	1.357
1993	5.760		**2.425**	0.718	**2.491**	**2.262**	0.764	**2.489**	**2.231**	1.516
1994	8.620		**2.892**	0.812	**2.942**	**2.706**	0.868	**2.939**	**2.635**	1.815
1995	8.350		**3.243**	0.860	**3.279**	**3.043**	0.923	**3.276**	**2.938**	2.047
1996	8.310		**3.406**	0.848	**3.425**	**3.211**	0.915	**3.422**	**3.068**	2.151
1997	8.290		**3.401**	0.793	**3.420**	**3.219**	0.860	**3.417**	**3.063**	2.215
1998	8.280		**3.337**	0.721	**3.353**	**3.178**	0.788	**3.350**	**3.004**	2.203
1999	8.280		**3.254**	0.644	**3.264**	**3.122**	0.709	**3.261**	**2.924**	2.133
2000	8.280		**3.245**	0.582	**3.260**	**3.132**	0.645	**3.257**	**2.920**	2.128
2001	8.280		**3.229**	0.513	**3.249**	**3.138**	0.572	**3.246**	**2.910**	2.115
2002	8.280		**3.186**	0.433	**3.212**	**3.119**	0.484	**3.209**	**2.877**	2.089
2003	8.280		**3.215**	0.358	**3.230**	**3.169**	0.399	**3.227**	**2.893**	2.114
2004	8.280		**3.358**	0.270	**3.365**	**3.335**	0.295	**3.362**	**3.014**	2.206
2005	8.194	3.450	**3.448**	0.092	**3.448**	**3.450**	0.000	**3.445**	**3.089**	2.228

Growth Rates Preserved SE: *(a)* 0.092 *(b)* 0.103 *(c)* 1.684 *(d)* Version 1

Table 5: **Predicted PPP Series for India**

YEAR	ER	ICP	UN	SE	UN-GRC *(a)*	CON	SE	No Reg-GRC *(b)*	No05-GRC *(c)*	PWT6.3
1971	7.490		**2.359**	0.378	**4.329**	**2.418**	0.449	**4.330**	**3.069**	2.887
1972	7.590		**2.483**	0.476	**4.605**	**2.541**	0.566	**4.607**	**3.265**	3.012
1973	7.740		**2.743**	0.523	**5.143**	**2.789**	0.613	**5.145**	**3.646**	3.342
1974	8.100		**2.914**	0.454	**5.505**	**2.935**	0.512	**5.506**	**3.902**	3.728
1975	8.380	2.594	**2.607**	0.145	**4.951**	**2.594**	0.000	**4.953**	**3.510**	3.371
1976	8.960		**2.691**	0.415	**4.961**	**2.706**	0.467	**4.963**	**3.517**	3.241
1977	8.740		**2.747**	0.505	**4.925**	**2.775**	0.584	**4.927**	**3.492**	3.213
1978	8.190		**2.701**	0.496	**4.718**	**2.729**	0.574	**4.720**	**3.345**	3.126
1979	8.130		**2.968**	0.451	**5.044**	**2.986**	0.506	**5.046**	**3.576**	3.190
1980	7.860	3.104	**3.116**	0.169	**5.156**	**3.104**	0.000	**5.158**	**3.655**	3.211
1981	8.660		**3.305**	0.488	**5.198**	**3.318**	0.546	**5.199**	**3.684**	3.295
1982	9.460		**3.538**	0.620	**5.278**	**3.569**	0.716	**5.279**	**3.741**	3.291
1983	10.100		**3.904**	0.674	**5.528**	**3.946**	0.779	**5.530**	**3.919**	3.490
1984	11.400		**4.246**	0.600	**5.723**	**4.287**	0.675	**5.725**	**4.057**	3.615
1985	12.400	4.667	**4.640**	0.235	**5.952**	**4.667**	0.000	**5.954**	**4.219**	3.739
1986	12.600		**4.845**	0.721	**6.217**	**4.880**	0.804	**6.219**	**4.407**	3.853
1987	13.000		**5.161**	1.012	**6.607**	**5.200**	1.161	**6.609**	**4.684**	3.988
1988	13.900		**5.429**	1.226	**6.918**	**5.465**	1.416	**6.920**	**4.904**	4.254
1989	16.200		**5.707**	1.404	**7.221**	**5.734**	1.625	**7.223**	**5.119**	4.427
1990	17.500		**6.126**	1.591	**7.685**	**6.137**	1.839	**7.688**	**5.448**	4.708
1991	22.700		**6.823**	1.840	**8.452**	**6.829**	2.127	**8.455**	**5.991**	5.141
1992	25.900		**7.365**	2.024	**8.992**	**7.366**	2.340	**8.995**	**6.375**	5.530
1993	30.500		**8.008**	2.211	**9.624**	**8.005**	2.554	**9.627**	**6.822**	5.849
1994	31.400		**8.748**	2.404	**10.338**	**8.734**	2.774	**10.341**	**7.328**	6.431
1995	32.400		**9.510**	2.571	**11.040**	**9.486**	2.964	**11.044**	**7.826**	6.960
1996	35.400		**10.193**	2.684	**11.618**	**10.169**	3.094	**11.622**	**8.236**	7.325
1997	36.300		**10.807**	2.751	**12.171**	**10.769**	3.167	**12.175**	**8.628**	7.604
1998	41.300		**11.689**	2.841	**12.987**	**11.651**	3.271	**12.991**	**9.206**	8.097
1999	43.100		**12.241**	2.804	**13.398**	**12.221**	3.230	**13.403**	**9.498**	8.370
2000	44.900		**12.555**	2.673	**13.574**	**12.528**	3.076	**13.579**	**9.623**	8.482
2001	47.200		**12.809**	2.476	**13.670**	**12.783**	2.843	**13.674**	**9.690**	8.603
2002	48.600		**13.254**	2.237	**13.957**	**13.239**	2.559	**13.962**	**9.894**	8.800
2003	46.600		**13.730**	1.923	**14.198**	**13.720**	2.183	**14.203**	**10.065**	9.076
2004	45.300		**14.206**	1.453	**14.439**	**14.212**	1.605	**14.444**	**10.236**	9.377
2005	44.272	14.670	**14.637**	0.502	**14.637**	**14.670**	0.000	**14.642**	**10.376**	9.540

Growth Rates Preserved SE: *(a)* 0.502 *(b)* 0.562 *(c)* 5.331.

Table 6: **Predicted PPP Series for Nigeria**

YEAR	ER	ICP	UN	SE	UN-GRC *(a)*	CON	SE	No Reg-GRC *(b)*	No05-GRC *(c)*	PWT6.3
1971	0.713		**0.332**	0.050	**0.503**	**0.349**	0.060	**0.504**	**0.553**	0.380
1972	0.658		**0.329**	0.062	**0.497**	**0.346**	0.076	**0.498**	**0.545**	0.373
1973	0.658		**0.328**	0.069	**0.496**	**0.345**	0.085	**0.497**	**0.544**	0.357
1974	0.630		**0.432**	0.096	**0.655**	**0.451**	0.116	**0.656**	**0.719**	0.388
1975	0.616		**0.487**	0.109	**0.739**	**0.503**	0.131	**0.740**	**0.811**	0.489
1976	0.627		**0.533**	0.117	**0.799**	**0.551**	0.140	**0.800**	**0.877**	0.571
1977	0.645		**0.561**	0.115	**0.831**	**0.579**	0.138	**0.833**	**0.913**	0.586
1978	0.635		**0.601**	0.108	**0.885**	**0.616**	0.127	**0.886**	**0.971**	0.619
1979	0.604		**0.622**	0.086	**0.911**	**0.631**	0.096	**0.912**	**1.000**	0.686
1980	0.547	0.643	**0.644**	0.030	**0.939**	**0.643**	0.000	**0.940**	**1.031**	0.754
1981	0.618		**0.688**	0.097	**0.997**	**0.689**	0.109	**0.999**	**1.095**	0.709
1982	0.673		**0.669**	0.117	**0.965**	**0.671**	0.134	**0.966**	**1.059**	0.753
1983	0.724		**0.753**	0.135	**1.078**	**0.755**	0.155	**1.079**	**1.183**	0.892
1984	0.767		**0.854**	0.127	**1.215**	**0.852**	0.142	**1.217**	**1.333**	1.025
1985	0.894	0.860	**0.866**	0.047	**1.222**	**0.860**	0.000	**1.224**	**1.341**	1.043
1986	1.750		**0.883**	0.140	**1.178**	**0.879**	0.155	**1.179**	**1.293**	1.301
1987	4.020		**1.367**	0.286	**1.720**	**1.361**	0.325	**1.722**	**1.888**	2.138
1988	4.540		**1.696**	0.402	**2.018**	**1.686**	0.461	**2.021**	**2.215**	2.584
1989	7.360		**2.492**	0.628	**2.807**	**2.474**	0.722	**2.812**	**3.082**	3.125
1990	8.040		**2.718**	0.700	**2.896**	**2.695**	0.805	**2.901**	**3.179**	3.420
1991	9.910		**3.349**	0.859	**3.363**	**3.331**	0.990	**3.368**	**3.692**	3.726
1992	17.300		**6.394**	1.579	**6.036**	**6.382**	1.824	**6.046**	**6.626**	5.999
1993	22.100		**10.161**	2.321	**9.006**	**10.182**	2.679	**9.021**	**9.886**	7.988
1994	22.000		**13.556**	2.695	**11.270**	**13.616**	3.091	**11.287**	**12.371**	11.283
1995	21.900		**22.192**	3.352	**17.226**	**22.390**	3.753	**17.253**	**18.909**	21.604
1996	21.900	32.539	**32.029**	1.645	**23.141**	**32.539**	0.000	**23.178**	**25.403**	26.434
1997	21.900		**30.142**	4.489	**23.071**	**30.297**	5.017	**23.107**	**25.325**	27.810
1998	21.900		**26.751**	5.147	**21.551**	**26.712**	5.879	**21.585**	**23.657**	30.520
1999	92.300		**28.309**	6.119	**23.854**	**28.187**	7.032	**23.891**	**26.185**	30.715
2000	102.000		**36.661**	8.296	**32.256**	**36.364**	9.532	**32.306**	**35.407**	32.090
2001	111.000		**38.177**	8.560	**34.879**	**37.819**	9.821	**34.934**	**38.287**	31.061
2002	121.000		**37.696**	7.897	**35.610**	**37.375**	9.030	**35.666**	**39.090**	37.348
2003	129.000		**43.642**	8.003	**42.167**	**43.313**	9.074	**42.233**	**46.288**	40.656
2004	133.000		**50.027**	6.995	**49.272**	**49.758**	7.722	**49.349**	**54.086**	47.085
2005	131.274	60.000	**60.096**	2.894	**60.096**	**60.000**	0.000	**60.190**	**65.968**	55.282

Growth Rates Preserved SE: *(a)* 2.894 *(b)* 3.240 *(c)* 30.916.

Table 7: **Predicted PPP Series for Honduras**

YEAR	ER	ICP	UN	SE	UN-GRC *(a)*	CON	SE	No Reg-GRC *(b)*	No05-GRC *(c)*	PWT6.3
1971	2.000		**0.861**	0.097	**1.274**	**0.910**	0.118	**1.201**	**1.306**	0.950
1972	2.000		**0.882**	0.126	**1.269**	**0.937**	0.156	**1.197**	**1.301**	0.943
1973	2.000		**0.901**	0.144	**1.266**	**0.958**	0.178	**1.194**	**1.298**	0.938
1974	2.000		**0.970**	0.163	**1.333**	**1.025**	0.201	**1.257**	**1.367**	0.984
1975	2.000		**0.961**	0.162	**1.296**	**1.005**	0.197	**1.222**	**1.328**	0.974
1976	2.000		**1.016**	0.166	**1.330**	**1.063**	0.201	**1.254**	**1.363**	0.932
1977	2.000		**1.098**	0.166	**1.403**	**1.143**	0.199	**1.323**	**1.438**	0.927
1978	2.000		**1.099**	0.145	**1.377**	**1.133**	0.171	**1.299**	**1.412**	0.928
1979	2.000		**1.148**	0.117	**1.418**	**1.167**	0.131	**1.337**	**1.453**	0.982
1980	2.000	1.202	**1.206**	0.042	**1.472**	**1.202**	0.000	**1.388**	**1.508**	1.005
1981	2.000		**1.229**	0.129	**1.442**	**1.252**	0.145	**1.359**	**1.478**	1.041
1982	2.000		**1.257**	0.181	**1.419**	**1.307**	0.213	**1.338**	**1.455**	1.041
1983	2.000		**1.341**	0.234	**1.460**	**1.419**	0.281	**1.377**	**1.497**	1.092
1984	2.000		**1.381**	0.273	**1.455**	**1.479**	0.333	**1.372**	**1.491**	1.078
1985	2.000		**1.453**	0.316	**1.485**	**1.571**	0.388	**1.401**	**1.523**	1.105
1986	2.000		**1.493**	0.353	**1.510**	**1.635**	0.439	**1.424**	**1.547**	1.077
1987	2.000		**1.509**	0.382	**1.511**	**1.671**	0.478	**1.424**	**1.549**	1.077
1988	2.000		**1.565**	0.418	**1.555**	**1.748**	0.527	**1.467**	**1.594**	1.069
1989	2.000		**1.626**	0.456	**1.604**	**1.827**	0.575	**1.513**	**1.645**	1.128
1990	4.110		**1.905**	0.558	**1.872**	**2.150**	0.705	**1.765**	**1.919**	1.327
1991	5.320		**2.328**	0.710	**2.280**	**2.642**	0.900	**2.150**	**2.337**	1.597
1992	5.500		**2.490**	0.788	**2.431**	**2.838**	1.001	**2.293**	**2.492**	1.714
1993	6.470		**2.771**	0.906	**2.700**	**3.171**	1.152	**2.546**	**2.768**	1.963
1994	8.410		**3.508**	1.187	**3.408**	**4.024**	1.510	**3.214**	**3.494**	2.436
1995	9.470		**4.304**	1.503	**4.171**	**4.950**	1.915	**3.933**	**4.276**	2.917
1996	11.700		**5.206**	1.874	**5.031**	**6.006**	2.390	**4.744**	**5.157**	3.625
1997	13.000		**6.229**	2.311	**6.051**	**7.191**	2.949	**5.706**	**6.203**	4.460
1998	13.400		**6.850**	2.618	**6.681**	**7.922**	3.345	**6.300**	**6.848**	4.968
1999	14.200		**7.510**	2.959	**7.347**	**8.709**	3.792	**6.928**	**7.531**	5.562
2000	14.800		**8.004**	3.252	**7.884**	**9.280**	4.173	**7.435**	**8.082**	5.941
2001	15.500		**8.384**	3.513	**8.312**	**9.721**	4.516	**7.838**	**8.520**	6.303
2002	16.400		**8.703**	3.760	**8.683**	**10.093**	4.848	**8.187**	**8.901**	6.582
2003	17.300		**9.176**	4.095	**9.162**	**10.633**	5.294	**8.639**	**9.392**	6.907
2004	18.200		**9.693**	4.466	**9.685**	**11.224**	5.794	**9.132**	**9.927**	7.249
2005	19.000		**10.337**	4.918	**10.337**	**11.966**	6.410	**9.747**	**10.596**	7.612

Growth Rates Preserved SE: *(a)* 4.918 *(b)* 5.461 *(c)* 5.052.

References

.. **, *.*.*. *, and *.*. *. *Modelling Spatially Correlated Error Structures in the Time-Space Extrapolation of Purchasing Power Parities*, volume Price Index Numbers in Time and Space of *Contributions to Statistics*, chapter Part II, pages 63–96. Springer-Verlag, Berlin, Germany, 2010.

*. *, *. *, and *. *. Extrapolation of purchasing power parities using multiple benchmarks and auxiliary information: A new approach. *Review of Income and Wealth*, page in press, 2010.

S. Ahmad. *Regression Estimates of per Capita GDP Based on Purchasing Power Parities*, volume International Comparisons of Prices, Output and Productivity: Contributions to Economic Analysis Series, chapter II, pages 237–264. Elsevier, 1996.

ADB Asian Development Bank. Purchasing power parities and real expenditures. 2005 international comparison program in asia and the pacific., 2007. URL `www.adb.org/Documents/Reports/ICP-Purchasing-Power-Expenditures/default.asp`.

R.J. Barro and X. Sala-i Martin. *Economic Growth*. MIT Press, Cambridge, MA, second edition, 2004.

J.H. Bergstrand. Structural determinants of real exchange rates and national price levels: Some empirical evidence. *American Economic Review*, 81(1):325–334, 1991.

J.H. Bergstrand. *Productivity, Factor Endowments, Military Expenditures, and National Price Levels*, volume International Comparisons of Prices, Output and Productivity: Contributions to Economic Analysis Series, chapter II, pages 297–317. Elsevier Science Publishers, 1996.

D. Chotikapanich, W.E. Griffiths, and D.S. Prasada Rao. Estimating and combining national income distributions using limited data. *Journal of Business and Economic Statistics*, 25: 97–109, 2007.

C. Clague. *Explanations of National Price Levels*, volume World Comparisons of Incomes, Prices and Product: Contributions to Economic Analysis Series, chapter III, pages 237–262. Elsevier, North Holland, 1988.

H.E. Doran. Constraining kalman filter and smoothing estimates to satisfy time-varying restrictions. *Review of Economics and Statistics*, 74(3):568–572, 1992.

J. Durbin and S.J. Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, 2001.

S.N. Durlauf, P. Johnson, and J. Temple. *Handbook of Economic Growth*, volume 1, chapter Growth Econometrics, pages 555–677. Elsevier, 2005.

A. C. Harvey and S. J. Koopman. Signal extraction and the formulation of unobserved components models. *Econometrics Journal*, 3:84–107, 2000.

A.C. Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University, Cambridge, 1989.

A.C. Harvey, T.M. Trimbur, and H.K. van Dijk. Trends and cycles in economic time series: A bayesian approach. *Journal of Econometrics*, 140(2):618–649, October 2007.

Alan Heston, Robert Summers, and B. Aten. Penn world table version 6.2, 2006. URL `http://pwt.econ.upenn.edu/php_site/pwt_index.php`.

G. Koop and H.K. van Dijk. Testing for integration using evolving trend and seasonal models: A bayesian approach. *Journal of Econometrics*, 97:261–91, 2000.

S. J. Koopman and A Harvey. Computing observation weights for signal extraction and filtering. *Journal of Economic Dynamics and Control*, 27:1317–1333, 2003.

I.B. Kravis and R.E. Lipsey. *Toward an Explanation of National Price Levels*. Princeton University Press, Princeton, 1983.

James LeSage and R. Kelly Pace. *Introduction ot Spatial Econometrics*, volume Statistics: Textbooks and Monographs. CRC Press. Taylor & Francis Group, 2009.

A. Maddison. *Monitoring the World Economy, 1820-1992*. Development Centre of the Organisation for Economic Co-operation and Development, Paris and Washington, D.C., 1995.

A. Maddison. *Contours of the World Economy 1-2030 AD: Essays in Macro-Economic History*. Oxford University Press, Oxford, 2007.

B. Milanovic. True world income distribution, 1988 and 1993: First calculation based on household surveys alone. *The Economic Journal*, 112:51–92, 2002.

M.H. Pesaran. General diagnostic tests for cross section dependence in panels. Technical report, CESifo Working Paper No. 1229. Category 10: Empirical and Theoretical Methods, 2004.

L. Pritchett. Divergence, big time. *Journal of Economic Perspectives*, 11:3–17, 1997.

D. Quah. Empirics for growth and distribution: Stratification, polarization, and convergence clubs. *Journal of Economic Growth*, 2:27–59, 1997.

A. Rose. Bilateral trade data set, January 2004. URL `http://faculty.haas.berkeley.edu/arose/RecRes.htm`.

X. Sala-i Martin. The world distribution of income: Falling poverty and ... convergence, period. *Quarterly Journal of Economics*, 121:351–97, 2006.

R. Summers and A. Heston. *Comparing the Income of Countries*, volume World Comparisons of Incomes, Prices and Product: Contributions to Economic Analysis Series, chapter III, pages 263–281. 1988.

R. Summers and A. Heston. The penn world table (mark 5): An expanded set of international comparisons, 1950-1988. *The Quarterly Journal of Economics*, 106(2):327–368, May 1991.

H. Theil. The development of international inequality, 1960-1985. *Journal of Econometrics*, 42:145–155, 1989.

UNDP. Statistics of the human development report, various. URL `http://hdr.undp.org/en/statistics/`.

A. The Invariance of the (Kalman Filter) Predictions to the Choice of Reference Country - Proposition 2.

A.1 Notation and Conventions

Without loss of generality we will take two reference countries as countries 1 and 2, and denote the $ln(PPP_t)$ relative to the two bases as $\boldsymbol{\alpha}_t^{(1)}$ and $\boldsymbol{\alpha}_t^{(2)}$. Other consequent notation will usually be obvious, making definition unnecessary.

By definition

$$\boldsymbol{\alpha}_t^{(2)} = \boldsymbol{\alpha}_t^{(1)} - \alpha_{2t}^{(1)} \tag{35}$$

Also $\alpha_{2t}^{(2)} \equiv \alpha_{1t}^{(1)} \equiv 0$.

Because the p_{it} is always zero for the base country, we will remove it from the Kalman filter cycle, and re-define $\boldsymbol{\alpha}_t^{(1)}$ and $\boldsymbol{\alpha}_t^{(2)}$ as the $N-1$ vectors $\boldsymbol{\alpha}_t^{(1)} = [\alpha_{2t}^{(1)}, \alpha_{3t}^{(1)}, ..., \alpha_{Nt}^{(1)}]'$ and $\boldsymbol{\alpha}_t^{(2)} = [\alpha_{1t}^{(2)}, \alpha_{3t}^{(2)}, ..., \alpha_{Nt}^{(2)}]'$. It follows from (35) that

$$\boldsymbol{\alpha}_t^{(2)} = \boldsymbol{A}\boldsymbol{\alpha}_t^{(1)} \tag{36}$$

where $\boldsymbol{A}$ is a non-stochastic, non-singular $(N-1) \times (N-1)$ matrix given by

$$\mathbf{A} = \begin{bmatrix} -1 & \mathbf{0}'_{N-2} \\ -\boldsymbol{j}_{N-2} & \boldsymbol{I}_{N-2} \end{bmatrix} \tag{37}$$

$\mathbf{j}_{N-2}$ is a vector of ones and $\mathbf{0}'_{N-2}$ a (row) vector of zeros.

Denoting the Kalman filter estimates obtained by using observations relative to the two base countries by $\hat{\boldsymbol{\alpha}}_t^{(1)}$and $\hat{\boldsymbol{\alpha}}_t^{(2)}$, the invariance property holds if it can be established that

$$\hat{\boldsymbol{\alpha}}_t^{(2)} = \boldsymbol{A}\hat{\boldsymbol{\alpha}}_t^{(1)} \tag{38}$$

A.2 Regression Estimates

a) *Benchmark years*

Estimates of β_{ot} and β_s are obtained by regressing benchmark observations $\tilde{\mathbf{p}}_t$ on the conditioning variables $\mathbf{x}_t = [x_{1t}, x_{2t}, ..., x_{N_1t}]'$ where we have taken countries $i = 1, 2, \ldots, N_1$, as the participating countries.

Now, by definition,

$$\tilde{\mathbf{p}}_t^{(2)} = \tilde{\mathbf{p}}_t^{(1)} - \tilde{p}_{2t}^{(1)}\mathbf{j}_{N-2} \tag{39}$$

That is, the dependent variable $\tilde{\mathbf{p}}_t^{(2)}$ is obtained by subtracting the same number $\tilde{p}_{2t}^{(1)}$ from each observation in $\tilde{\mathbf{p}}_t^{(1)}$. Because the regressors $\mathbf{X}_t$ do not change when the base country is changed from 1 to 2, by standard regression theory

$$\hat{\beta}_{0t}^{(2)} = \hat{\beta}_{0t}^{(1)} - \tilde{p}_{2t}^{(1)} \tag{40}$$

$$\hat{\boldsymbol{\beta}}_s^{(2)} = \hat{\boldsymbol{\beta}}_s^{(1)} = \hat{\boldsymbol{\beta}}_s$$

That is, intercepts change but slopes are invariant. It follows that for non-participating countries

$$\hat{\mathbf{p}}_t^{(2)} = \hat{\mathbf{p}}_t^{(1)} - \hat{p}_{2t}^{(1)}\mathbf{j}$$

Thus, defining the "observation vector" $\mathbf{y}_t$ by $\mathbf{y}_t = [\tilde{\mathbf{p}}_t, \hat{\mathbf{p}}_t]'$ and discarding the base country observation (as it is always zero) we have

$$\mathbf{y}_t^{(2)} = \mathbf{A}\mathbf{y}_t^{(1)} \tag{41}$$

b) *Non-benchmark years*

Here the observation is the regression prediction $\hat{\beta}_o^{(i)}\mathbf{j}_N + \mathbf{X}_t\hat{\boldsymbol{\beta}}_s$ $(i = 1, 2)$. We now adjust the observation by subtracting the base country prediction from all predictions. This ensures the base country observation is zero, and the value of the intercept is irrelevant.

Then,

$$y_{it}^{(1)} = (x'_{it} - x'_{1t})\hat{\boldsymbol{\beta}}_s; \qquad y_{it}^{(2)} = (x'_{it} - x'_{2t})\hat{\boldsymbol{\beta}}_s$$

$$(x'_{it} - x'_{1t})\hat{\beta}_s - (x'_{2t} - x'_{1t})\hat{\boldsymbol{\beta}}_s = y_{it}^{(1)} - y_{2t}^{(1)}$$

Thus,

$$\mathbf{y}_t^{(2)} = \mathbf{A}\mathbf{y}_t^{(1)} \tag{42}$$

It follows from (41) and (42) that for both benchmark and non-benchmark years, the fundamental transformation $\mathbf{y}_t^{(2)} = \mathbf{A}\mathbf{y}_t^{(1)}$ holds.

A.3 The covariance of the measurement error

The measurement error in the benchmark *PPPs* and growth rates are assumed to have a covariance proportional to the form[30]:

$$\mathbf{V}_t = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \sigma_{1t}^2 \boldsymbol{j}\boldsymbol{j}' + diag(\sigma_{2t}^2, ..., \sigma_{Nt}^2) \end{bmatrix} \tag{43}$$

where, σ_{it}^2 is the variance of country i at time t and σ_{1t}^2 is the variance of the reference country.

Let $\mathbf{V}_t^{(1)}$ the $(N-1) \times (N-1)$ matrix obtained by ignoring the first row and column of $\mathbf{V}_t$,

$$\mathbf{V}_t^{(1)} = \sigma_\eta^2[\sigma_{1t}^2\mathbf{j}\mathbf{j}' + diag(\sigma_{2t}^2, ..., \sigma_{Nt}^2)]$$

Then,

$$\mathbf{A}\mathbf{V}_t^{(1)}\mathbf{A}' = \sigma_\eta^2[\sigma_{2t}^2\mathbf{j}\mathbf{j}' + diag(\sigma_{1t}^2, \sigma_{3t}^2, ..., \sigma_{Nt}^2)] = \mathbf{V}_t^{(2)}$$

[30]See RRD for a formal derivation of $\mathbf{V}_t$ from the definition of PPP.

A.4 The observation equation

The fundamental observation equation used in the method is

$\mathbf{y}_t = \boldsymbol{\alpha}_t + \boldsymbol{\zeta}_t, \qquad E(\boldsymbol{\zeta}_t\boldsymbol{\zeta}_t') = \mathbf{H}_t$

where $\mathbf{y}_t$ is an observation of the unobserved state; $\boldsymbol{\alpha}_t$ and $\boldsymbol{\zeta}_t$ is an observation error.

Because $\boldsymbol{\alpha}_t^{(2)} = \mathbf{A}\boldsymbol{\alpha}_t^{(1)}$ by definition and $\mathbf{y}_t^{(2)} = \mathbf{A}\mathbf{y}_t^{(1)}$ by regression properties and construction (see previous sections) it follows that $\boldsymbol{\zeta}_t^{(2)} = \mathbf{A}\boldsymbol{\zeta}_t^{(1)}$. And thus because $\boldsymbol{A}$ is non-stochastic,

$$\mathbf{H}_t^{(2)} = \mathbf{A}\mathbf{H}_t^{(1)}\mathbf{A}' \tag{44}$$

This is the fundamental result that enables us to prove invariance.

A.5 The transition equation

The transition equation used is of the form

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \boldsymbol{\eta}_t \tag{45}$$

where,

c_t is the observed growth rate of $\boldsymbol{\alpha}_t$; $\boldsymbol{\eta}_t$ is an error with $E(\boldsymbol{\eta}_t) = 0$ and $E\left(\boldsymbol{\eta}_t\boldsymbol{\eta}_t'\right) \equiv \mathbf{Q}_t{=}\sigma_\eta^2\mathbf{V}_t$.

By defining $\mathbf{V}_t$ as in (43), it follows that,

$$\mathbf{Q}_t^{(2)} = \mathbf{A}\mathbf{Q}_t^{(1)}\mathbf{A}' \tag{46}$$

A.6 Proposition 2 Proof

For the reader's reference the Kalman filter equations, are given by

Prediction Equations

$\hat{\boldsymbol{\alpha}}_{t|t-1} = \hat{\boldsymbol{\alpha}}_{t-1} + \mathbf{c}_t$; $\mathbf{P}_{t|t-1} = \mathbf{P}_{t-1} + \mathbf{Q}_t$

Updating Equations

$\mathbf{F}_t = \mathbf{P}_{t|t-1} + \mathbf{H}_t$; $\hat{\boldsymbol{\alpha}}_t = \hat{\boldsymbol{\alpha}}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{F}_t^{-1}(\mathbf{y}_t - \hat{\boldsymbol{\alpha}}_{t|t-1})$; $\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{F}_t^{-1}\mathbf{P}_{t|t-1}$

Assume,

$$\hat{\boldsymbol{\alpha}}_{t-1}^{(2)} = \mathbf{A}\hat{\boldsymbol{\alpha}}_{t-1}^{(1)} \tag{47}$$

from which it follows (because $\boldsymbol{A}$ is non-stochastic) that

$$\mathbf{P}_{t-1}^{(2)} = \mathbf{A}\mathbf{P}_{t-1}^{(1)}\mathbf{A}' \tag{48}$$

Following the Kalman filter covariance cycle

$$\mathbf{P}_{t|t-1}^{(2)} = \mathbf{P}_{t-1}^{(2)} + \mathbf{Q}_t^{(2)} = \mathbf{A}\mathbf{P}_{t-1}^{(1)}\mathbf{A}' + \mathbf{A}\mathbf{Q}_t^{(1)}\mathbf{A}' \text{ (by(46))}$$

$$\mathbf{P}_{t|t-1}^{(2)} = \mathbf{A}\mathbf{P}_{t|t-1}^{(1)}\mathbf{A}' \tag{49}$$

$$\mathbf{F}_t^{(2)} = \mathbf{P}_{t|t-1}^{(2)} + \mathbf{H}_{t|t-1}^{(2)} = \mathbf{A}\mathbf{P}_{t|t-1}^{(1)}\mathbf{A}' + \mathbf{A}\mathbf{H}_t^{(1)}\mathbf{A}' \text{ (by (43))}$$

$$\mathbf{F}_t^{(2)} = \mathbf{A}\mathbf{F}_t^{(1)}\mathbf{A}' \tag{50}$$

The updating equation for $\hat{\boldsymbol{\alpha}}_t^{(2)}$ is

$$\hat{\boldsymbol{\alpha}}_t^{(2)} = \hat{\boldsymbol{\alpha}}_{t-1}^{(2)} + \mathbf{P}_{t|t-1}^{(2)}\left(\mathbf{F}_t^{(2)}\right)^{-1}(\mathbf{y}_t^{(2)} - \mathbf{c}_t^{(2)} - \hat{\boldsymbol{\alpha}}_{t-1}^{(2)})$$

Substituting using 44, 47, 48 and 41,

$$\hat{\boldsymbol{\alpha}}_t^{(2)} = \mathbf{A}\hat{\boldsymbol{\alpha}}_{t-1}^{(1)} + \mathbf{A}\mathbf{P}_{t|t-1}^{(1)}\mathbf{A}'\left(\mathbf{A}\mathbf{F}_t^{(1)}\mathbf{A}'\right)^{-1}(\mathbf{A}\mathbf{y}_t^{(1)} - \mathbf{A}\mathbf{c}_t^{(1)} - \mathbf{A}\hat{\boldsymbol{\alpha}}_{t-1}^{(1)})$$

$$= \mathbf{A}[\hat{\boldsymbol{\alpha}}_{t-1}^{(1)} + \mathbf{P}_{t|t-1}^{(1)}\left(\mathbf{F}_t^{(1)}\right)^{-1}(\mathbf{y}_t^{(1)} - \mathbf{c}_t^{(1)} - \hat{\boldsymbol{\alpha}}_{t-1}^{(1)})] \text{ (because } \boldsymbol{A} \text{is non-singular)} \tag{51}$$

From the definition of $\boldsymbol{c}_t$ following equation (7), it is clear that $\boldsymbol{c}_t^{(2)} = \boldsymbol{A}\boldsymbol{c}_t^{(1)}$.

Thus,

$$\hat{\boldsymbol{\alpha}}_t^{(2)} = \mathbf{A}\hat{\boldsymbol{\alpha}}_t^{(1)} \tag{52}$$

It follows by induction that if the estimation is commenced when (52) holds, invariance will be true for all subsequent years.

B. Preserving Movements in Implicit GDP Deflators through the Smoothing Filter

In this appendix we show that using a fixed interval smoother with $\sigma_\eta^2 = 0$, the resulting smoothed estimates of the state vector, $\mathbf{a}_{t|\mathrm{T}}$, preserve the movement in the implicit price deflator and the covariance matrix of the smoothed estimate equals the Kalman filter estimate of the covariance at time T for all t.

The equations of a fixed interval smoother are,

$$\mathbf{a}_{t|T} = \hat{\boldsymbol{\alpha}}_t + \mathbf{P}_t^*(\mathbf{a}_{t+1|T} - \mathbf{c}_{t+1} - \hat{\boldsymbol{\alpha}}_t) \tag{53}$$

$$\mathbf{P}_{t|T} = \mathbf{P}_t + \mathbf{P}_t^*(\mathbf{P}_{t+1|T} - \mathbf{P}_{t+1|t})\mathbf{P}_t^{*'} \tag{54}$$

$$\mathbf{P}_t^* = \mathbf{P}_t\mathbf{P}_{t+1|t}^{-1} \tag{55}$$

where, $\hat{\boldsymbol{\alpha}}_t$ is the Kalman filter estimate of the state vector; $\mathbf{P}_t$ is the Kalman filter unconditional covariance of the state vector ; $\mathbf{P}_{t+1|t}$ is the Kalman filter conditional covariance of the state vector; $\mathbf{a}_{t|T}$ is the Kalman smoothed estimate of the state vector; $\mathbf{P}_{t|T}$ is the covariance of $\mathbf{a}_{t|T}$.

Now, if $\sigma_\eta^2 = 0$, $\mathbf{P}_{t+1|t} = \mathbf{P}_t$, which from (55) implies $\mathbf{P}_t^* = \mathbf{I}_N$. Therefore, $\mathbf{a}_{t|\mathrm{T}} = \mathbf{a}_{t+1|\mathrm{T}} - \mathbf{c}_{t+1}$, or $\mathbf{a}_{t+1|\mathrm{T}} = \mathbf{a}_{t|\mathrm{T}} + \mathbf{c}_{t+1}$. That is, smoothed estimates, $\mathbf{a}_{t|\mathrm{T}}$ preserve the movement in the implicit price deflator.

Now consider the covariance matrix in (54). Since, $\mathbf{P}_{t+1|t} = \mathbf{P}_t$ and $\mathbf{P}_t^* = \mathbf{I}_N$ we have, $\mathbf{P}_{t|\mathrm{T}} = \mathbf{P}_{t+1|\mathrm{T}}$. Thus, $\mathbf{P}_{t|\mathrm{T}}$ is constant with respect to t and $\mathbf{P}_{t|\mathrm{T}} = \mathbf{P}_{T|T} = \mathbf{P}_{T|T}$ for all t.