Division of Economics
A.J. Palumbo School of Business Administration and
McAnulty College of Liberal Arts
Duquesne University
Pittsburgh, Pennsylvania

Aggregation Bias in the US Personal Consumption Expenditure

Steve Bronder

Submitted to the Economics Faculty

In partial fulfillment of the requirements for the degree of

Bachelor of Science in Business Administration

Fall 2014

Abstract

The purpose of this analysis is to test for aggregation bias in six levels of cross-sectional aggregation in the United States Personal Consumption Expenditure. Aggregation bias can cause aggregate inflation data to overstate evidence of inflation divergence. Whether or not inflation rates diverge has policy implications such as changes to target inflation. This paper uses first and second generation panel unit root tests on the National Income and Product Accounts that make up the PCE. Aggregation bias exists if NIPA inflation rates converge or diverge at different levels of aggregation.

Using the second generation panel unit root test developed by Bai and Ng (2004, 2010), this study finds aggregation bias within the PCE aggregates. These results are consistent with theoretical and empirical literature supporting that improper aggregation of indices can lead to erroneously concluding inflation rates non-convergent. Possible policy implications include changing the way the aggregate is constructed by giving heavier weight to more persistent series.

JEL Classification: E40, C43, C01, C50

Keywords: Applications, Aggregation, PCE, Macroeconometrics

Contents

1 Introduction

The purpose of this analysis is to test for aggregation bias in the United States Personal Consumption Expenditure (PCE). This paper uses first and second generation panel unit root tests on the National Income and Product Accounts (NIPA) that make up the PCE. Second generation tests differ from first generation tests in that the latter drop the assumption of cross sectional independence in the error term[1]. Higher levels of aggregation are made to represent the lower, more dis-aggregate levels. Aggregation bias exists if the aggregate data and underlying data do not share the same convergence process. Aggregation is important because the process used to aggregate the data may remove information and create divergent aggregate inflation rates when dis-aggregate inflation rates converge. Monetary policy of the Federal Open Market Committee (FOMC) is based on a target inflation rate, however there are concerns that if the FOMC focuses on aggregate inflation it may cause individual sectors to diverge. Clark (2006) uses dis-aggregate quarterly NIPA accounts to study the distribution of inflation persistence across consumption sectors. Inflation persistence is the tendency of inflation to revert slowly to its equilibrium or long run level after a shock. Clark's results indicate that the average across consumption sectors for the persistence of dis-aggregate inflation is below aggregate persistence. When Clark accounts for a mean break, average dis-aggregate persistence is found to be similar to aggregate inflation persistence.

Clark uses a sum of auto-regressive coefficients model, similar to Levin and Piger (2003). The sum of auto-regressive coefficients uses the sum of the vector auto-regressions for each series at a specified lag to determine a single set of lagged coefficients for the entire panel. However, Dias and Carlos (2010) find that the sum of auto-regressive coefficients may be subject to significant biases coming from the presence of outliers in the data or model misspecification. The sum of auto-regressive coefficients model may

[1] For more information see Hurlin (2007).

also be subject to significant bias from the presence of additive outliers, which can be caused by a large number of regressors such as those found in dis-aggregate inflation data. Pesaran and Smith (1995) show that pooling and aggregating coefficients from dynamic heterogeneous panels across groups can give inconsistent and possibly misleading estimates of coefficient values. This is due to coefficients not differing in a random manner and biasing so that they do not represent the underlying coefficients.

The dynamic properties of the data can be affected by how the aggregates are constructed and can cause aggregate levels of inflation to not actually represent underlying dis-aggregate data. Higher levels of aggregation can appear less stationary due to trends in the underlying data becoming obfuscated. While previous studies such as Byrne and Fiess (2010) and Dreger and Kosfeld (2010) have examined idiosyncratic and common components, they did not perform the new tests created in Bai and Ng (2010). Bai and Ng (2004, 2010) use the factor structures[2] of large dimensional panels to detect if non-stationarity in a panel is pervasive, variable specific, or both.

This paper examines whether aggregate inflation measures inflate evidence of inflation divergence. Common effects can become more pervasive and less stationary due to the process of aggregation affecting the dynamic properties of the data. If aggregation bias exists it may be wise to alter the aggregation method used or change weights for each sector. This study is not attempting to determine the cause of Personal Consumer Expenditure inflation rates or how much inflation rates differ for aggregate and dis-aggregate data. Unless it is possible to dismiss aggregation bias, examining the determinants of inflation rates is secondary. This paper adds to the current aggregation bias literature because of its use of the NIPA underlying tables that were updated in

[2]Large dimensional panels can be expressed by a smaller number of unobserved variables known as factors. These are derived from linear orthogonal projections of the scaled data which explain the largest variance of the data.

2009 and the methodology used to test for aggregation bias. The Panel Analysis of Nonstationarity in the Idiosyncratic and Common Components (PANIC) methodology accounts for cross-sectional correlation within the factors and error term that previous studies have not accounted for. This paper also implements two new tests from Bai and Ng (2010) and measures the power of these tests through a Markov Chain Monte Carlo method.

The second generation test used in this paper is the PANIC approach developed by Bai and Ng (2004,2010). The number of common factors are estimated by the BIC_3 and AIC_3[3] found in Bai and Ng (2002). When using a factor model on data with possible cross-sectional correlation, less common factors are needed to describe the variability in the data than other heuristic methods would prescribe. These information criteria provide a data driven method to estimate the number of common factors. Because of this, PANIC is a powerful approach because it accounts for cross-sectional correlation in the factors and residual as well as a multiple factor structure. This is beneficial to the study of aggregation bias as each of these anomalies can bias the results.

2 Inflation Convergence

This paper uses a time series approach to convergence created in Bernard and Durlauf (2005). Convergence is defined for two sectors i and j by:

$$\lim_{k \to \inf} E(\pi_{it+k} - \pi_{jt+k} | I_t) = 0 \quad (1)$$

[3]The formulas for these are $BIC_3 = V(k, \hat{F}^k) + k\hat{\sigma}^2\left(\frac{(N+T-k)ln(NT)}{NT}\right)$ and $AIC_3 = V(k, \hat{F}^k) + k\hat{\sigma}^2\left(2\frac{(N+T-k)}{NT}\right)$ where N and T are the number of cross-sections and time, respectively. $V(k, \hat{F}^k)$ is the average residual variance when k factors are assumed for each cross-section. $\hat{F}^k$ is the estimator of the factor scores. $\hat{\sigma}^2$ is a consistent estimate of the sums of the expected value of the residual variance across sectors and time periods divided by NT.

Where π is the inflation rate per sector i, j at time t given an I_t co-integration process. While this definition of convergence is normally used for country-by-country inflation convergence, this paper will use the sectors of NIPA accounts. When differences between sectors converge to zero, convergence is absolute. If the inflation rate converges to a constant, convergence is conditional.

Previous literature that considers inflation convergence uses panel unit root approaches. Kocenda and Papell (1997) and Holmes (2002) use data from the European Union (EU) and show convergence of national inflation rates until the 1990s. This makes sense as during that time countries were attempting to conform to the standards necessary to join the EU. Busetti et al. (2007) use inflation differentials in the EU from 1980 to 2004 and test subsets of the data from 1980-1997 and 1998-2004. When using a multivariate Dicky-Fuller they reach the same conclusion as Kocenda and Papell (2007). However, when testing the subset from 1998-2004 using a multivariate version of the Kwiatkowski et al. (1992) stationarity test they reject convergence.

There are three statistical anomalies that may bias an analysis on inflation convergence. These include cross-sectional correlation, structural breaks, and aggregation bias. First generation unit root tests, such as Im et al. (2003) and Hadri (2000), make the assumption that each residual is independent of one another. By not accounting for cross-sectional correlation a size distortion is found which creates a tendency to reject the null hypothesis when the null hypothesis is valid.

Structural breaks refer to shifts in the mean of inflation in one sector introducing breaks in the convergence process of a series. This creates a bias in favor of divergence. Angeloni et al. (2006) test sectoral inflation rates for six European Union countries before the creation of the monetary union and after. They find an increase in inflation persistence which could be due to structural changes in private inflationary expectations.

Aggregation biases time series analysis towards rejecting inflation convergence. Altissimo et al. (2006, 2009) state that the process of aggregation plays an important role in defining the characteristics of the aggregate series. Heterogeneity across sectors is likely to lead to bias in aggregation as the aggregate shifts improperly due to the variances in each series. Their study gives strong evidence of persistence of inflation differentials. Paya et al. (2007) test the PCE for temporal aggregation bias from yearly to quarterly to monthly using several scalar measures of persistence including the sum of auto-regressive coefficients, largest auto-regressive root, and half life. The largest auto-regressive root method uses the largest root of an auto-regressive model to determine how long the effects of persistence will exist. The half life measures how long it takes an auto-regressive process to revert back halfway to its mean. She finds that aggregation bias is induced when aggregating from month, to quarter, to year. While Paya et al. (2007) tests for temporal aggregation bias, this paper tests whether the PCE contains aggregation bias when aggregating across sectors.

This paper uses an analysis similar to the one Byrne and Fiess (2010) use to test aggregation bias and convergence of inflation differentials in the EU. Byrne and Fiess (2010) use the Harmonized Index of Consumer Prices[4] gathered by Eurostat for thirteen European Union countries and run first and second generation unit root and stationarity tests. Byrne and Fiess's findings suggest aggregation bias within the HICP aggregates. While aggregate EU national inflation rates are diverging, dis-aggregate inflation rates are converging. They conclude that aggregation appears to bias evidence in favor of

[4]The HICP is an indicator of inflation and price stability developed by the European Central Bank. It is a consumer price index built by a weighted average of price indices for member states who have adopted the euro.

divergence.

Similar to Clark (2006), this paper uses NIPA data from the Bureau of Economic Analysis to test whether inflation rate are relevant to the level of aggregation. Six levels of aggregation come from the underlying tables found on the BEA's website. The highest aggregate contains goods and services. The data dis-aggregates into 201 series such as furniture and household equipment, books and maps, magazines, newspapers, and sheet music. The analysis tests the hypothesis that inflation rates diverge at higher levels of aggregation, but converge at lower, dis-aggregate levels. The data is monthly price indices from 1959M1 to 2014M8 ($T = 657$) for data dis-aggregates of up to 201 sectors. This paper uses year on year PCE inflation rate ($\pi_t = 100 \times \ln(p_t/p_{t-12})$).

3 Econometric Tests

This section reviews statistical tests for examining non-stationarity in inflation rates for each level of aggregation. Some panel methods such as the Hadri (2000), Im et al. (2003), and Levin et al (2002) account for cross-sectional heterogeneity, but are based on the assumption that the residuals for each cross-section are independent. The test from Bai and Ng (2004) drop the assumption of cross-sectional independence in the residual of first generation tests which allows for valid pooling methods for tests on the residual. Three additional tests, the Panel Modified Sargan–Bhargava (PMSB) and Moon and Perron (MP), and a bias corrected PANIC test on the idiosyncratic component (P_a, P_b), found in Bai and Ng (2010), are used to analyze unit roots in the idiosyncratic component. MP estimates an unbiased pooled auto-regressive coefficient and PMSB uses a sample moment. Using a sample moment allows the test to account for unknown population parameters by estimating the population moments for short, long, and one sided variance.

3.1 Panel ADF Tests

The first test is the Im et al. (2003) panel version of the augmented Dickey-Fuller (ADF) test. This model tests if inflation rates are stationary and comes from the following panel auto-regressive model:

$$\pi_{it} = \alpha_i + (\phi_i + 1)\pi_{it-1} + \epsilon_{it} \mid i = 1, ..., N \; ; t = 1, ..., T \tag{2}$$

This model includes α_i, a heterogeneous constant, $(\phi_i + 1)$, a heterogeneous auto-regressive parameter, and ϵ_{it}, which is assumed to be a cross-sectionally independent residual. Transform equation (2) to a panel regression that investigates the change in inflation rates as follows:

$$\Delta\pi_{it} = \alpha_i + \phi_i \Delta\pi_{it-1} + \epsilon_{it} \tag{3}$$

Im et al. (2003) has a null hypothesis that the process for all time series in equation (3) are random walks with drift.

$$H_0 : \phi_1 = \phi_2 = ... = \phi_N = \phi = 0 \tag{4}$$

And a heterogeneous alternative hypothesis of

$$H_A : \phi_1 < 0, ...\phi_{N_1} < 0, |N_1 < N \tag{5}$$

This alternative hypothesis differs from the test created in Levin et al. (2003) in that the auto-regressive parameter (ϕ_i) does not have to be stationary for all cross sections. This means that heterogenous alternative hypothesis used in Im et al. (2003) is less restrictive. Accepting the null hypothesis of a unit root in each series implies non-convergence in the aggregate. The test statistic (Z_{IPS}) for Im et al. (2003) is defined as the following:

$$Z_{IPS} = \frac{\sqrt{N}[\bar{\tau} - E(\tau_i)]}{\sqrt{Var(\tau_i)}} \rightarrow N(0,1) \tag{6}$$

Using central limit theory, the test statistic is distributed standard normal for t_i, the test statistic of the individual ADF tests. The first moment correction $E(\tau_i)$ and second moment correction $Var(\tau_i)$ are corrections that center and scale the test statistic.

$$\bar{\tau} = \frac{1}{N \sum_{i=1}^{N} t_i} \tag{7}$$

3.2 Stationarity as The Null Hypothesis

This paper implements the test for non-stationarity found in Hadri (2000), which uses a null hypothesis of no unit root in any series versus an alternative of a unit root existing in the panel. This null is useful because it allows a test of non-divergence. The Hadri test statistic is based on the residuals of an OLS regression:

$$\pi_{it} = \mu_{it} + \epsilon_{it} \tag{8}$$

$$\mu_{it} = \mu_{it-1} + u_{it} \tag{9}$$

$$H_0 : \ \sigma_u^2 = 0 \tag{10}$$

The inflation rate (π_{it}) are regressed on a constant (μ_{it}). The ϵ_{it} and u_{it} are independent and normally distributed error terms with variance σ_ϵ^2 and σ_u^2. If the variance of u_{it} is equal to zero then μ_{it} becomes a constant and the inflation rate is stationary. The Hadri test statistics (Z_μ) is made from the averages of the univariate KPSS test found in Kwiatkowski et al. (1992). The test is distributed standard normal as follows:

$$Z_\mu = \frac{\sqrt{N}(L\bar{M}_\mu)}{\varsigma_\mu} \rightarrow N(0,1) \tag{11}$$

The average of the individual KPSS test is defined in Hadri (2000) as the following:

$$L\bar{M}_\mu = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\frac{1}{T^2}\sum_{t=1}^{N} S_{it}^2}{\hat{\sigma}_{\epsilon i}^2}\right) \tag{12}$$

Where $\hat{\sigma}_{\epsilon i}^2$ is an estimator of the variance of the error term in the time series being observed and S_{it} is the partial sum of the residuals ϵ_{it} in eq 8. Busetti et al. (2007) suggest a test for convergence should use a null hypothesis of stationarity when it is possible that inflation rates are similar for each series. This paper's analysis will focus more on unit root tests because they can provide evidence that shocks to inflation rates are permanent.

3.3 Panel Analysis of Non-stationarity in Idiosyncratic and Common Components

The PANIC approach introduced by Bai and Ng (2004) uses the factor structure inherent in large panels to investigate non-stationarity. The data generating process is the follows:

$$\pi_{it} = D_i + \lambda_i' F_t + e_{it} \tag{13}$$

$$F_t = \alpha F_{t-1} + u_t \tag{14}$$

$$e_{it} = \rho_{i0} e_{it-1} + \epsilon_{it} \tag{15}$$

The series π_{it} is the sum of a deterministic component (D_i), a common component $\lambda_i' F_t$ and an error e_{it}, the idiosyncratic component. The deterministic component can be 0, a fixed effect, or 1, a fixed effect and trend. Inflation rates diverge if the common factor $(\alpha = 1)$ or the idiosyncratic component $(\rho_{i0} = 1)$ are non-stationary. Gengenbach et al. (2004) shows that the common factors and idiosyncratic component can be analyzed individually because they are independent of on another. This allows for the null of non-convergence for each of the PANIC based tests to be used in both parts of the factor model. The correct number of factors is determined by the information criteria procedure developed in Bai and Ng (2002). The panel Akaike information criterion is used for the three highest levels of aggregation as Bai and Ng (2002) shows that the BIC_3 can underestimate the number of common components when N is small. However, Bryne and Fiess (2010) point out that the BIC_3 is more robust to cross-sectional correlation.

This paper uses an ADF test on the common factor ($ADF^c_{\hat{F}}$) and a pooled ADF test on the idiosyncratic individual errors ($ADF^c_{\hat{e}}(i)$) that is developed from fisher's method of pooling p-values. Each of these come from Bai and Ng (2004). A fisher pooling method is an aggregation method for p-values with the same null hypothesis and assumes independence of p-values. The idiosyncratic component's test statistic is distributed as standard normal as given by (16)

$$P^c_{\hat{e}} = \frac{-2\sum_{i=1}^{N} \log p(i) - 2N}{\sqrt{4N}} \to N(0,1) \tag{16}$$

The ADF test on the idiosyncratic component has a p-value $p(i)$ for each i cross-section. The test statistic tests whether $H_0 : \rho_i = 1 \forall i$ against $H_A : \exists\, i$ where $\rho_i < 1$

This paper also implements two additional tests from Bai and Ng (2010). The PMSB is based off of the modified Sargan-Bhargava test and uses a sample moment while the Moon and Perron (2004) (MP) test estimates the pooled auto-regressive coefficient for its test statistic. In order to replace unknown population moments, a sample moment is found by deriving the moment generating process for the k-th moment, drawing a sample, and then estimating the population moment from the sample. The test statistics requires the short-run, long-run, and $\hat{\phi}_\epsilon^4$ of ϵ_{it} defined as the following.

$$\sigma_{\epsilon i}^2 = E(\epsilon_{it}^2) = \sum_{j=0}^{\infty} \rho_{ip}^2, \;\; \omega_{\epsilon i}^2 = \big(\sum_{j=0}^{\infty} \rho_{ip}\big)^2, \;\; \hat{\phi}_\epsilon^4 = \frac{1}{N}\sum_{i=1}^{N}(\hat{\omega}_{\epsilon i}^2)^2 \tag{17}$$

The PMSB test statistic is defined as the following.

$$PMSB = \frac{\sqrt{N}(tr(\frac{1}{NT^2}\hat{e}'\hat{e}) - \hat{\omega}_\epsilon^2/6)}{\sqrt{\hat{\phi}_\epsilon^{\,4}/45}} \tag{18}$$

The trace [tr()] of an n-by-n square matrix is the sum of the elements on the main diagonal, the diagonal from the upper left to the lower right. For the MP test, the auto-regressive component ρ can also be estimated from the data by the following formula.

$$X_{it} = (1-\rho L)D_{it} + \rho X_{it-1} + u_{it}, \; u_{it} = \lambda_i' f_t + \epsilon_{it} \tag{19}$$

Where L is the lag operator. With formula eighteen Bai and Ng (2010) create an incidental trend model C based on Moon and Perron (2004) that has a fixed effect and trend $D_{it} = a_i + b_{it}$ This model defines two MP test statistics from Bai and Ng (2010).

$$t_a = \frac{\sqrt{N}T(\hat{\rho}^+ - 1)}{\sqrt{K_a \hat{\phi}_\epsilon^4 / \hat{\omega}_\epsilon^4}} \tag{20}$$

$$t_b = \sqrt{N}T(\hat{\rho}^+ - 1)\sqrt{\frac{1}{NT^2} tr(X'_{-1} M_z X_{-1}) K_b \frac{\hat{\omega_\epsilon^2}}{\hat{\phi_\epsilon^4}}} \tag{21}$$

The bias-corrected, defactored, pooled OLS estimator $\hat{\rho}^+$ is available in Moon and Perron (2004). When the data is demeaned and detrended (Model C), M_z =

$I - z(z'z)^{-1}z'$, $z = (1,t)'$, $\hat{\phi}_\epsilon = -\hat{\sigma}^2_\epsilon/2$, $K_a = 15/4$, and $K_b = 4$. The parameter z is built to demean and detrend the data.

The analysis also implements a bias-corrected version of the pooled test on the idiosyncratic component of Bai and Ng (2004). P_a and P_b are based on a bias-corrected pooled PANIC auto-regressive parameter ρ and a model in which the deterministic component detrends the data.

$$\hat{\rho}^+ = \frac{tr(\hat{e}'_{-1}\hat{e})}{tr(\hat{e}'_{-1}\hat{e}_{-1})} + \frac{3}{T}\frac{\hat{\sigma}^2_\epsilon}{\hat{\omega}^2_\epsilon} \tag{22}$$

$$= \hat{\rho} + \frac{3}{T}\frac{\hat{\sigma}^2_\epsilon}{\hat{\omega}^2_\epsilon} \tag{23}$$

The left side of the addition on the right side is the original auto-regressive parameter with the bias-correction on the right. The bias-correction makes use of the short run over long run variance and the number of observations. The test statistics are the following.

$$P_a = \frac{\sqrt{N}T(\hat{\rho}^+ - 1)}{\sqrt{(36/5)\hat{\phi}^4_\epsilon\hat{\sigma}^4_\epsilon/\hat{\omega}^8_\epsilon}} \tag{24}$$

$$P_b = \sqrt{N}T(\hat{\rho}^+ - 1)\sqrt{\frac{1}{NT^2}(tr(\hat{e}'_{-1}\hat{e}_{-1})\frac{5}{6}\frac{\hat{\omega}^6_\epsilon}{\hat{\sigma}^4_\epsilon\hat{\sigma}^4_\epsilon}} \tag{25}$$

The parameters $\hat{\sigma}^2_\epsilon, \hat{\omega}^2_\epsilon, \hat{\phi}^4_\epsilon$ are defined in (17).

4 Results

4.1 First Generation Panel Unit Root Test

This section gives results for the panel unit root and non-stationarity tests. Table one gives results for the Im et al. (2003) panel unit root and the Hadri (2000) panel stationarity tests. The Im et al. (2003) Z_{ips} panel unit root test rejects the null hypothesis for all six levels of aggregation. Evidence of convergence is present for all six levels of aggregation. However, the Hadri (2000) Z_{μ} also rejects the null hypothesis, implying that at least one of the panels contains a unit root. The first generation tests suggest some inflation rates in the panels are converging while some are diverging.

Table 1: First generation panel unit root and stationarity tests

	Aggregate	Im et al.	Hadri
1	Aggregate 1	-4.17*	78.26*
2	Aggregate 2	-4.67*	138.29*
3	Aggregate 3	-11.70*	221.21*
4	Aggregate 4	-30.68*	351.61*
5	Aggregate 5	-49.24*	518.23*
6	Aggregate 6	-57.54*	580.32*

Asterisk (*) denotes rejection of the null hypothesis of unit root. Alternative hypothesis are as follows; Hadri: At least one series has unit roots, Im: Stationary.

The Im et al. panel unit root test is not robust to large N panel sets. O' Connel (1998) shows that large N panels can potentially cause size distortion effects and will lead to a tendency to reject the null hypothesis. Giulietti (2009) finds the panel variant of the KPSS test developed by Hadri (2000) suffers from size distortions in the presence of cross section dependence. This leads the robustness of the test statistics to be in serious question. The PANIC method is shown in Bai and Ng (2004, 2010) to be robust for large N x T matrices and so will be used to further test the PCE.

4.2 Second Generation Panel Unit Root Test

Panel unit roots are tested by using Bai and Ng's (2004, 2010) PANIC approach with the number of common factors decided by the information criteria found in Bai and Ng (2002). This model is more robust than the previous tests because it tests for non-stationarity in the idiosyncratic component, the common component, or both. One of the main improvements the PANIC method brings is that the common components are data determined and not heuristically decided.

Table two contains the results for the tests on the common and idiosyncratic component. For aggregate one, PANIC identifies two common factors and shows different properties for the idiosyncratic and common components. The null hypothesis of a unit root in sectoral inflation rates for each level of aggregation is rejected and so the idiosyncratic components for each level of aggregation is found to be stationary. These results are similar to the results Bryne and Fiess (2010) find for the idiosyncratic component in the Harmonized Index of Consumer Prices of the European Union.

Table 2: PANIC Tests

	Aggregate	$P^c_{\hat{e}}$	Factors	$ADF^c_{\hat{F}}$					
1	Aggregate 1	4.14*	2	-4.2*	-2.72				
2	Aggregate 2	7.34*	4	-4.11*	-4.87*	-3.1*	-2.69		
3	Aggregate 3	17.21*	6	-7.4*	-6.58*	-3.61*	-3.13*	-6.15*	-5.77*
4	Aggregate 4	32.98*	2	-6.74*	-9.02*				
5	Aggregate 5	57.27*	3	-12.66*	-6.6*	-7.048*			
6	Aggregate 6	70.47*	4	-12.56*	-6.55*	-6.77*	-7.63*		

Asterisk (*) denotes rejection of the null hypothesis of unit root. 5% critical value of -2.86 for the factor unit root test $ADF^c_{\hat{F}}$. The idiosyncratic unit root test, $P^c_{\hat{e}}$ has a 5% critical value of 1.64. Lag lengths are determined by the formula $4(T/100)^{\frac{1}{4}}$ as done in Bai and Ng (2004). The dimensions of the aggregates are (N=2,T=657),(N=4,T=657),(N=17,T=657),(N=68,T=657),(N=152,T=657),(N=201,T=657), respectively

The information criterion from Bai and Ng (2002) identify two and four common components for aggregates one and two, respectively. Because each of these aggregates has at least one common component that fails to reject the null hypothesis there is evidence to suggest non-convergence in aggregates one and two. In aggregates three through six there are six, two, three, and four common components, respectively. For these levels of aggregation this analysis rejects the null hypothesis for each common component and conclude aggregates three through six converge. These results imply aggregation bias within the Personal Consumption Expenditure.

This paper implements the PMSB, P_a, P_b, and Model C for the MP test from Bai and Ng (2010) in order to investigate the presence of a unit root within the idiosyncratic Component. If the tests on the idiosyncratic show higher aggregates have non-convergence and lower aggregates converge then there is evidence of aggregation bias. Table three holds these results.

Table 3: PANIC Tests on Idiosyncratic Component

				Model C	
Aggregates	PMSB	P_a	P_b	t_a	t_b
Aggregate 1	1.018	.405	.055	-0.359	-0.211
Aggregate 2	1.370	.727	1.000	-0.249	-0.186
Aggregate 3	-2.629*	-8.186*	-4.678*	-1.137	-0.731
Aggregate 4	-4.462*	-19.982*	-9.523*	-7.011*	-3.071*
Aggregate 5	-6.010*	-33.644*	-14.180*	-5.056*	-2.148*
Aggregate 6	-6.610*	-45.342*	-17.476*	-4.187*	-1.741*

Asterisk (*) denotes rejection of the null hypothesis at the .05 level. Has a critical value a -1.64. Alternative hypothesis is stationarity.

The PMSB, P_a, and P_b tests both reject the null hypothesis for aggregates three through six. Model C rejects for aggregates four through six. This is reasonable evidence to conclude that aggregates three through six converge. Aggregates one and two do not reject the null hypothesis and so these aggregates are deemed non-convergent. Model C does not reject for aggregate three, however this may be due to the power of the test statistic being weak for small N and T structures. While these results imply aggregation bias, it is important to investigate the power of the test statistics.

5 Power Tests

Due to the extreme difference in the number of sectors in each aggregate this paper will conduct a Markov Chain Monte Carlo experiment in order to test against size distortions. Samples are generated from the posterior distribution of a normal theory factor analysis model. The factor scores F_t, and factor loadings λ_i' have normal priors. ρ_i of equation thirteen holds an Inverse Gamma prior with values of .5 in the diagonal. The inverse gamma prior is the conjugate that allows the factor loadings and scores to have a normal prior. Since the posterior will be the same distribution as the prior this conjugate is necessary for the normal factor model. Gibbs Sampling is then performed to simulate from the posterior joint distribution of the factors still decided by Bai and Ng (2002). Gibbs sampling refers to the process of obtaining a sequence of random samples from a probability distribution. The algorithm samples variables from the joint probability distribution conditional on the last values of the other variables[5]. The posterior marginal densities are assumed to approximate the distribution of the test statistic which makes it possible to infer directly the probability of rejecting the null hypothesis given the possibility of the null being false.

For each level 100,000 chains are burned before sampling another 100,000 while thinning every 10. Performing the Geweke (1992) test for convergence and checking for non-significant autocorrelation show that the chain has reasonably converged for these values. The importance of this process is that once the chain has hit a convergent point PANIC is run over each segment in order to find the posterior marginal distribution of each test statistic. Tables four and five contain the results of the power tests for each

[5]For more information see Casella (1992)

level of aggregation.

Table 4: PANIC 2004 Power Results

Aggregate	$P_{\hat{e}}^{c}$	Factors	$ADF_{\hat{F}}^{c}$					
Aggregate 1	0.7372	2	0.6076	0.3725				
Aggregate 2	0.8787	4	0.5645	0.5842	0.5876	0.7696		
Aggregate 3	1.00	6	0.5667	0.5657	0.5483	0.5579	0.5407	0.5792
Aggregate 4	1.00	2	1.00	0.9707				
Aggregate 5	1.00	3	1.00	0.9943	1.00			
Aggregate 6	1.00	4	1.00	0.9205	1.00	1.00		

The results in tables five and six confirm the Monte Carlo results in Bai and Ng (2004, 2010), but for the six samples in this paper. For the tests in Bai and Ng (2004), the size of the tests on the common component are nominal, however the tests on the idiosyncratic component appear overpowered. The power tests for PMSB, P_a, P_b, and Model C are consistent with the power results from PANIC (2010). Each test has nominal power properties except for model C which has worse nominal power properties compared to the other models. The PMSB test appears to have the best nominal power properties while Model C and P_b appear to have the worst power properties. There is sufficient evidence to conclude size distortions are highly unlikely to explain the convergence behavior of aggregate and dis-aggregate levels.

Table 5: PANIC 2010 Power Tests on Idiosyncratic Component

				Model C	
Aggregates	PMSB Tests	P_a	P_b	t_a	t_b
Aggregate 1	0.1472	0.0251	0.0075	0.0077	0
Aggregate 2	0.4209	0.3013	0.0108	0.0067	0.0041
Aggregate 3	0.4126	0.2656	0.0379	0.0053	0.0372
Aggregate 4	0.943	0.9023	0.9993	0.9591	0.8443
Aggregate 5	1.00	1.00	1.00	1.00	1.00
Aggregate 6	1.00	1.00	0.9996	0.9865	1.00

6 Conclusion

The purpose of this analysis is to test for aggregation bias in six levels of aggregation in the United States Personal Consumption Expenditure. This paper uses first and second generation panel unit root tests on the National Income and Product Accounts that make up the PCE. Aggregation bias exists if different levels of aggregation do not have the same convergence process. While previous studies have examined the PCE for aggregation bias, no papers have used the NIPA tables to examine aggregation bias since they were updated in 2009 to include lower levels of dis-aggregation. This paper uses panel tests that are robust to cross-sectional correlation and an ad-hoc power study to check the robustness of the results.

Using the second generation panel unit root test developed by Bai and Ng (2004, 2010), this study finds aggregation bias between the highest two levels of aggregation and the lowest four levels. The results of this analysis finds that aggregation introduces non-stationary common factors. These common factors are not present in the dis-aggregate data, which cause aggregates one and two to diverge. These results are theoretically and empirically consistent with the research of Pesaran and Smith (2005), Imbs et al (2005a, b), Altissimo et al (2009), and Byrne and Fiess (2010) supporting that improper aggregation of indices can lead to erroneously concluding inflation rates non-convergent.

The monetary policy of the Federal Reserve targets national inflation. There are concerns that if the Federal Reserve uses policy based on aggregate inflation it may lead to divergence of inflation rates for individual sectors which could threaten the long term recovery of the united states economy. Policy implications related to the persistence of inflation rates range from support for fiscal stimulus policy to no action and even to raise target inflation in order to avoid deflation in some sectors of the economy. One policy suggestion from this analysis is that it may be advisable to change how the aggregate is constructed. In the presence of aggregation bias, placing more weight on sectors with persistent dynamics may help the Federal Reserve's ability to achieve stable inflation.

References

[1] Altissimo, Filippo, Laurent Bilke, Andrew Levin, Thomas Matha, and Benoit Mojon. "Sectoral and Aggregate Inflation Dynamics in the Euro Area." Journal of the European Economic Association 4.2-3 (2006): 585-593. Print.

[2] Andrew D. Martin, Kevin M. Quinn, Jong Hee Park. MCMCpack: Markov Chain Monte Carlo in R. Journal of Statistical Software. 42(9) (2011): 1-21.URL: http://www.jstatsoft.org/v42/i09/.

[3] Casella, George, and Edward I. George. "Explaining the Gibbs Sampler." The American Statistician 46.3 (1992): 167-74. Print.

[4] Christian Dreger & Reinhold Kosfeld, 2007. "Do Regional Price Levels Converge?: Panel Econometric Evidence Based on German Districts,"Discussion Papers of DIW Berlin 754, DIW Berlin, German Institute for Economic Research

[5] Christophe Hurlin, Valerie Mignon. Second Generation Panel Unit Root Tests. 2007.<halshs-00159842>

[6] Giulietti, Monica, Jesús Otero, and Jeremy Smith. "Testing for Stationarity in Heterogeneous Panel Data in the Presence of Cross-Section Dependence." Journal of Statistical Computation and Simulation 79.2 (2009): 195-203. Print.

[7] Stock, James H. "A Class of Tests for Integration and Cointegration. Mimeo, Departement of Economics, Harvard University.

[8] Stock, James H., and Mark W. Watson. "Testing for Common Trends." Journal of the American Statistical Association 83.404 (1988): 1097. Print.

[9] Yves Croissant, Giovanni Millo. Panel Data Econometrics in R: The plm Package. Journal of Statistical Software 27(2) (2008). URL http://www.jstatsoft.org/v27/i02/.

[10] Altissimo, Filippo, Benoit Mojon, and Paolo Zaffaroni. "Can Aggregation Explain the Persistence of Inflation?." Journal of Monetary Economics 56.2 (2009): 231-241. Print.

[11] Angeloni, Ignazio, Luc Aucremanne, and Matteo Ciccarelli. "Price Setting and Inflation Persistence: did EMU Matter?." Economic Policy 21.46 (2006): 353-387. Print.

[12] Bai, Jushan, and Serena Ng. "A PANIC Attack on Unit Roots and Cointegration." Econometrica 72.4 (2004): 1127-1177. Print.

[13] Bai, Jushan, and Serena Ng. "Determining the Number of Factors in Approximate Factor Models." Econometrica 70.1 (2002): 191-221. Print.

[14] Bernard, Andrew B., and Steven N. Durlauf. "Convergence in International Output." Journal of Applied Econometrics 10.2 (1995): 97-108. Print.

[15] Blanchard, Olivier Jean, Robert J. Gordon, and Christopher A. Sims. "Aggregate and Individual Price Adjustment." Brookings Papers on Economic Activity 1987.1 (1987): 57. Print.

[16] Byrne, Joseph P and Norbert Fiess. "Euro Area Inflaton: Aggregation Bias and Convergence." Review of World Economics 146, no.2 (2010): 339-357.

[17] Cecchetti, Stephen G., Nelson C. Mark, and Robert J. Sonora. "Price Index Convergence Among United States Cities*." International Economic Review 43.4 (2002): 1081-1099. Print.

[18] Chen, Shiu-Sheng, and Charles Engel. "Does Aggregation Bias Explain The ppp Puzzle?." Pacific Economic Review 10.1 (2005): 49-72. Print.

[19] Clark, Todd E.. "Disaggregate Evidence on the Persistence of Consumer Price Inflation." Journal of Applied Econometrics 21.5 (2006): 563-587. Print.

[20] Dias, Daniel A., and Carlos Robalo Marques. "Using Mean Reversion as a Measure of Persistence." Economic Modelling 27.1 (2010): 262-273. Print.

[21] E., Kocenda, and Papell D.H.. "Inflation Convergence Within the European Union: A Panel Data Analysis." International Journal of Financial Economics 2.3 (1997): 189-198. Print.

[22] F., Busetti, Forni L., Harvey A., and Venditti F.. "Inflation Convergence and Divergence Within the European Monetary Union." International Journal of Central Banking 2.3 (2007): 95-121. Print.

[23] Gengenbach, Christian, Franz C. Palm, and Jean-Pierre Urbain. "Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling." Econometric Reviews 29.2 (2009): 111-145. Print.

[24] Geweke, J. "Evaluating the accuracy of sampling-based approaches to calculating posterior moments." Bayesian Statistics 4. Clarendon Press, Oxford, UK.

[25] Hadri, Kaddour. "Testing for Stationarity in Heterogeneous Panel Data." The Econometrics Journal 3.2 (2000): 148-161. Print.

[26] Holmes, Mark J.. "Panel data Evidence on Inflation Convergence in the European Union." Applied Economics Letters 9.3 (2002): 155-158. Print.

[27] Imbs, Jean, Haroon Mumtaz, Morten O. Ravn, and Helene Rey. "PPP Strikes Back: Aggregation and the Real Exchange Rate*." Quarterly Journal of Economics 120.1 (2005): 1-43. Print.

[28] K, Im, Pesaran M. H., and Shin Y. "Testing For Unit Roots In Heterogeneous Panels." Journal of Econometrics 115.1 (2003): 53-74. Print.

[29] Kwiatkowski, Denis, Peter C.b. Phillips, Peter Schmidt, and Yongcheol Shin. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root." Journal of Econometrics 54.1-3 (1992): 159-178. Print.

[30] Levin, A., C.-F. Lin, and C.-S. J. Chu. 2002. "Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties". Journal of Econometrics 108: 1–24.

[31] Levin, Andrew , and Jeremy Piger. "Is inflation Persistence Intrinsic in Industrial Economies." European Central Bank Working Paper Series 0334 (2004): n. pag. European Central Bank Series of Working Papers. Web. 15 Sept. 2014.

[32] Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu. "Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties." Journal of Econometrics 108.1 (2002): 1-24. Print.

[33] M.H., Pesaran , and Smith R.P.. "Estimating Long-run Relationships from Dynamic Heterogeneous Panel." Journal of Econometrics 68.1 (1995): 79-113. Print.

[34] M.H., Pesaran. "A Simple Panel Unit Root Test in the Presence of Cross-sectional Dependence." Journal of Applied Econometrics 22.2 (2007): 265-312. Print.

[35] Moon, Hyungsik Roger, and Benoit Perron. "Testing for a Unit Root in Panels with Dynamic Factors." Journal of Econometrics 122.1 (2004): 81-126. Print.

[36] O' Connel, P.G.. "The Overvaluation of Purchasing Power Parity." Journal of International Economics 44.1 (1998): 1-19. Print.

[37] Paya, Ivan, Agustin Duarte, and Ken Holden. "On the Relationship between Inflation Persistence and Temporal Aggregation." Journal of Money, Credit and Banking 39.6 (2007): 1521-1531. Print.

[38] Pesaran, M. Hashem, Ron P. Smith, Takashi Yamagata, and Lyudmyla Hvozdyk. "Pairwise Tests of Purchasing Power Parity." Econometric Reviews 28.6 (2009): 495-521. Print.

[39] U.S. Bureau of Economic Analysis, "Table 2.4.4U. Prince Indexes for Personal Consumption Expenditure by Type of Product," (accessed November 1, 2014).