BBVA RESEARCH

Nº 16/10

Working Paper

MAY 2016

Alfonso Ugarte Ruiz

Long run and short run components in explanatory variables and differences in Panel Data estimators

Long run and short run components in explanatory variables and differences in Panel Data estimators

Alfonso Ugarte Ruiz[1,2]

May 2016

Abstract

We investigate the idea that when we separate an explanatory variable into its "between" and "within" variations we could be roughly decomposing it into a structural (long-term) and a cyclical component respectively, and this could translate into different Between and Within estimates in panel data. We first present empirical evidence that the long and medium-term components of the GDP per capita have indeed different effects than the cyclical (short-term) component on the private Credit-to-GDP ratio. Then, we show through Monte Carlo simulations that the former explanation fits better the empirical results than the alternative based on a misspecification of the dynamics in order to explain the observed differences in the Between and Within estimators.

Keywords: Between estimator, Within estimator, cyclical, structural, short-term component, long-term component, Credit-to-GDP ratio, permanent income hypothesis.

JEL: C01, C18, C23, C33, C51, C58, G20, G21.

1: This paper has been accepted for presentation at the "22nd International Conference of Panel Data" that this year will take place in Curtin Business School in Perth, Australia, and at the "New Trends and Developments in Econometrics" Conference organized by Banco de Portugal. We thank Georghi Kolev for helpful comments and suggestions and seminar participants in the 2015 Pompeu Fabra PhD Meeting, the same as all the participants in our BBVA-Research internal seminars.
2: BBVA Economics Research Department, Madrid. Email: alfonso.ugarte@bbva.com. The views expressed are those of the authors and do not necessarily reflect those of BBVA-Research or BBVA.

1 Introduction and Motivation

The first studies that pooled and compared time-series and cross-section estimates found that they may differ widely when estimating the same underlying equation. The initial interpretation was that cross-sectional estimators reflect long-run responses while time-series studies tend to yield short-run responses (Kuh (1959) and Houthakker (1965)).

In applied work, the Between and the Within estimators still tend to differ very often, although the initial interpretation seems to be long-forgotten based on the argument that the observed differences in the estimators are only due to different misspecification problems. For instance, Mundlak (1978) qualifies the theoretical difference in the Between and Within estimators as "imaginary" and based only on the correlation between the unobservable omitted individual effects and the explanatory variables.

Later on, Baltagi and Griffin (1984) took side with Mundlak although because of a different rationale. They claim that the difference exists but that it is due to a different misspecification problem. They argue that it is the consequence of estimating a static specification when the true underlying data comes from a dynamic process that includes a number of lags of the explanatory variables. Several other theoretical studies have confirmed that in a dynamic framework, the Between estimator should capture the "long-term" effect and that the Within estimator will tend to capture a short-term (or contemporaneous) effect, although with different biases depending on the dynamic characteristics of the explanatory variables (Van den Doel and Kiviet (1994, 1995), Pirotte (1999), Eggen and Pfaffermayr (2005)).

Other authors (Griliches and Hausman (1986), Mairesse (1990), Mairesse and Sassenou (1991) have argued that the differences between the cross-sectional and time-series estimators could arise from measurement errors in the explanatory variables that could be minimized by the Between estimator.

In any case the initial idea that the estimators could be different in nature because they truly capture different types of effects seems to have been almost completely abandoned despite the potential implications of such concept for the panel data theory.

In this paper we want to somehow revisit the old interpretation about the difference in the Between and Within estimators but from a different perspective. We propose that in certain situations, when we separate the variance of an explanatory variable into its "between" and "within" parts we could also be roughly decomposing it into a structural or long-term component and a temporary or cyclical component respectively, and therefore, this could translate into different Between and Within estimates.

This idea is not completely new. For instance, in the reference manual of the program Stata for the command "xtreg", it is suggested that "*changes in the average value of x (an explanatory variable) for an individual may have a different effect from temporary departures from the average. In an economic situation, y might be purchases of some item and x income; a change in average income should have more effect than a transitory change*"[3]. Despite the simplicity and appeal of this idea, we were not able to find a theoretical or empirical literature supporting this argumentation.

In this paper we further explore such concept from both a theoretical and an empirical point of view. In order to support this idea we first investigate the effects of real income per capita on the private credit ratio (Credit-to-GDP), as an empirical example that provides evidence that the observed difference in the Between and the Within estimators could be due to actual different effects of the cyclical and structural components of the explanatory variable, i.e. GDP per capita.

3: Stata 12 Longitudinal Data/Panel Data Reference Manual, Release 12.

The relationship between these two variables offers a perfect example, because the permanent income hypothesis suggests a theoretically different relation between the two variables given temporary vs. permanent changes in income. Long-run variations (permanent) in income should have a positive impact but short-run changes (temporary) should have a negative impact.

Actually, when we estimate the empirical relationship between the private Credit-to-GDP ratio and real GDP per capita we find the following results:

1. The Between and the Within estimators are clearly different in any specification according to the standard Hausman test. This would initially imply that we should hold to the Within estimated effect according to the standard textbook suggestion.
2. Moreover, when taking into account the effect of time fixed-effects, the Between estimator is clearly *higher* than the Within one.
3. However, when we decompose GDP per capita into three components, a long-term (individual) mean, a medium-term component, and a cyclical component, we clearly find that the effect of the cyclical component is *negative*, as the permanent income hypothesis suggests. On the other hand, the long-term and the medium-term components' effects turn out to be *positive*, also in accordance to the expected theoretical results.
4. Similarly, when we estimate the effect of the medium and short-term components through the Within or the Arellano-Bond estimators, we find that the effect of the short-term component is *negative* and the medium-term one is *positive*, again, in accordance to the expected sign suggested by the economic theory.

The fourth result is quite important because it shows that we can get a *different* estimated effect of two components (short- vs. medium-term) while using estimators that are *consistent* to the presence of both time and individual effects (Within and Arellano-Bond). Moreover, we also obtain the same result through an estimator such as the Arellano-Bond that explicitly accounts for a dynamic specification of the underlying data. Thus, *the estimated differences cannot be explained by the existing theoretical explanations*[4].

Through Monte Carlo simulations we also show that the empirical results explained before are, on the one hand, consistent with our proposed explanation of different effects of the different components and that on the other hand, they could not be properly explained by the alternative explanation based on a dynamic misspecification. In the simulations we consider two types of misspecifications.

In the first simulation the estimated regression specification omits that the true data generating process includes three components with their own different effects (components misspecification), with true *positive* long- and medium-term effects and a true *negative* short-term one. In the second simulation the estimated specification ignores that true data generating process includes multiple lags for the explanatory variables (dynamic misspecification). This leads to the following testable predictions:

1. When the true long-term effect is positive in a dynamic misspecification, the bias of the Within estimator is always positive no matter the size of the sample[5].
2. The simulated Within coefficient estimate decrease when we reduce the length of the time series (T) in both type of misspecifications, but the reduction is much higher in the components misspecification (when the true short-term component effect is negative).

4: We have also tested for possible non-stationarity of our dependent and explanatory variable. However, the different panel unit-root tests for the Credit-to-GDP ratio and the GDP per capita (in logs, PPP and real USD) seem to provide mixed and inconclusive results at best, although in the majority of the cases non-stationarity is rejected for both variables.
5: Although the size of the bias does depend on T.

3. On the one hand, the simulated Arellano-Bond estimate is not sensitive to the length of the sample in the case of the dynamic misspecification. On the other hand, under the components misspecification, the Arellano-Bond estimator is indeed sensible to the number of periods. If the true medium-term effect is positive and the true short-term effect is negative, the Arellano-Bond estimate changes from a positive (or near zero) to a negative value when T becomes small.

The previous theoretical simulation results are further confirmed in the empirical data:

- The first prediction is not consistent with our empirical finding of a *positive* Between estimate and a *negative* short-term component effect. In order to find a negative short-term response in the dynamic misspecification case the true long-term response must be negative as well (given the positive bias).
- The Within estimate strongly decreases when we estimate it in small T subsamples (3 to 5 years long) compared to the original sample from 1990-2014 (25 years), becoming non-significant and close to zero on average, and even negative in many of the sub-samples.
- The Between estimate does not change in small T subsamples (on average).
- The Arellano-Bond estimate is positive when estimated in the whole sample (25 years), but when estimated in small T subsamples (3 to 5 years) it actually becomes negative both on average and in the majority of the subsamples.

The implications of these results are both theoretical and practical. First, the presented rationale for the difference in the estimated effects could help us reconcile some conflicting empirical results. Moreover, it implies that in the situations in which the Between and Within estimators could be different for the reasons exposed before, we may have to re-evaluate the use of some specification tests widely used in panel data analysis, such as the Hausman test and its implied choice between the Within and Between estimator.

The rest of the paper is organized as following: Section 2 discusses the possible theories behind the empirical differences in the Within and the Between estimators and offers a set of testable hypotheses derived from our proposed explanation of this phenomenon. Section 3 presents some empirical estimations of the relationship between private credit (leverage) and real income per capita. Section 4 presents the results of the Monte Carlo simulations that compare which theoretical explanation fits better the empirical results shown in Section 3. Section 5 concludes.

2 Theoretical Discussion

The typical error component, panel-data model specification is the following:

$$Y_{it} = \beta X_{it} + u_{it} \tag{1}$$

$$u_{it} = \delta_i + \gamma_t + \varepsilon_{it} \tag{2}$$

We usually estimate one of the following equations or a combination of the two (as in Random Effects):

$$\bar{Y}_{i\cdot} = \beta^B \bar{X}_{i\cdot} + \bar{u}_{i\cdot} \tag{3}$$

$$(Y_{it} - \bar{Y}_{i\cdot}) = \beta^W (X_{it} - \bar{X}_{i\cdot}) + (u_{it} - \bar{u}_{i\cdot}) \tag{4}$$

We generally assume that the Between and the Within coefficients are identical, i.e. $\beta^B = \beta^W$.

However, as explained in the introduction, since the first econometric studies started pooling time-series and cross-section data it has been frequently common to find that these estimators differ substantially when estimating the same underlying equation. For instance, Kuh (1959) provided evidence of such difference for an investment equation. Similarly, Houthakker (1965) through an exercise on consumers' expenditure in different countries found that the effect estimates from "within countries" and "between countries" regressions were significantly different both statistically and economically. They concluded that this difference arises because cross-section (*Between*) estimators reflect long-run effects, whereas time series deviations (*Within*) reflect shorter run reactions.

However, Mundlak (1978) argued that the observed differences in the estimators are only due to the omission of individual fixed effects that are correlated with the explanatory variables. He thus argued that there is a unique estimator, that there should be a uniform approach, and that the Within estimator is the best one, because is consistent since it eliminates the bias from the unobserved individual effects.

Later on, Baltagi and Griffin (1984) and others have argued that the observed differences were only the consequence of estimating (1) when the true underlying data comes from any of the following dynamic processes:

$$Y_{it} = \sum_{r=0}^{k} \beta_r X_{it-r} + u_{it} \tag{5}$$

$$Y_{it} = \sum_{r=1}^{k} \rho_r Y_{it-r} + \beta X_{it} + u_{it} \tag{6}$$

$$X_{it} = \sum_{r=1}^{k} \lambda_r X_{it-r} + \nu_{it} \tag{7}$$

These studies generally conclude that the Between estimator captures well the true long-run effect, whereas the Within estimator underestimates the long-run, although it approximates better the short-run response. However, there could be different biases depending on the dynamic structure of the explanatory and dependent variables and the size and length of the sample.

For instance, Van den Doel and Kiviet (1994) conclude that “static estimators usually underestimate the long run effect”, if the explanatory variable is stationary. They show that “the estimators are consistent for the long run effect” if the explanatory variable follows a random walk and if the impact of the initial condition is negligible. Van den Doel and Kiviet (1995) conclude that “the long run coefficients will usually be underestimated when the Fixed Effects “Within” estimator is employed”.

Pirotte (1999) assumes a general, ADL(l,*k*) dynamic error components model. He estimates that the within estimator tends to approximate the short run effects if the data generating process of the explanatory variable exhibits no memory, whatever the memory of the dependent variable. He furthers estimates that the Between estimator of the static model converges, in all cases, to long run effects and that long run effects are obtained directly from the static relation without the need of a dynamic model.

Eggen and Pfaffermayr (2005) provide estimates of the biases of the within, between, random effects and pooled OLS with respect to both the short and long run effects of a dynamic data generating process in the case one wrongly estimates a static model. An important result they find and that we will use later is that given a positive true short-run parameter, the within estimate is upward (downward) biased, if the memory of the explanatory variable is high (low). At a high memory of the explanatory variable, the Within estimator gets even closer to the long-run than to the short run impact, and at small *T*, its bias with respect to the short run increase as *T* rises.

Another reason why the initial idea of different estimators could have been left aside is the success of dynamic panel data estimators such as the Arellano-Bond, which provides a proper specification for equations like (6) and in some cases to specifications like (5)[6]. These dynamic estimators could estimate appropriately the short-run response β in both equations (5) or (6), and the long-run response would be simply given by $\beta/(1-\rho)$, where ρ is the autocorrelation coefficient of the lagged dependent variable.

Other authors (Griliches and Hausman (1986), Mairesse (1990), Mairesse and Sassenou (1991) have argued that the differences between the cross-sectional and time-series estimators could arise from measurement errors in the explanatory variables. The Between estimator tends to minimize the importance of those errors by averaging, meanwhile the Within estimator magnifies the variability of these measurement errors and increases the resulting bias. Mairesse and Sassenou (1991) also described other reasons for the persistent differences found in empirical studies that analyze the effect of R&D in productivity. However, most of the reasons exposed have a practical rather than theoretical rationale, and are mostly applied to their particular case of study.

Summarizing, we can broadly identify the following reasons in the literature for the divergence between the Within and the Between estimators:

1. Inconsistency of the Between estimator due to unobserved individual (or time) fixed effects.
2. Misspecification of the true dynamic underlying process. The Within estimator captures the contemporaneous effect βX_{it}, meanwhile the Between estimator captures the accumulated effect of all the lags or dynamic components ($\sum_{r=0}^{k} \beta_r X_{it-r}$ or approximately $\beta X_{it}/(1-\rho)$).
3. Measurement error in the explanatory variables.

However, we want to consider the case in which the explanatory variables have different cyclical and structural components. For instance, if some underlying determinants of the explanatory variables are constant, others vary slowly over time and others vary widely and in a noisy way.

Consider for instance the case in which the GDP or the GDP per capita are the explanatory variable. We would all agree that the GDP could be composed of a cyclical component and a structural component. For instance, we recognize that the observed differences in the GDP per capita level between different countries could depend on a number of institutional, regulatory or geographical factors that do not change or change slowly over time, but also on assets or commodity prices that change very fast.

This would not matter if all the different components of GDP or GDP per capita have the same effect on the dependent variable we want to explain. But what if different components of income have different effects on the dependent variable? For instance, if the dependent variable responds to an economic decision of agents, they should take into consideration what the factors behind movements or differences in income are.

6: In the cases in which the effect of each lag of the explanatory variable decays as a function of the autocorrelation coefficient of the dependent variable in the following way: $\beta_r = \rho^r \beta$

For instance, when agents take decisions on investment, or when they decide about granting or demanding credit, they should consider whether their current income level responds only to transitory factors or to more permanent and stable causes.

The main idea that we want to consider is that in many economic situations, when we separate the variance of an explanatory variable into its “between” and “within” parts we are roughly decomposing it into a structural or long-term component and a temporary or cyclical component respectively, and therefore, this could translate into different Between and Within estimates.

The specific example we will consider here is the relationship between private leverage (or financial deepening) and economic development (income per capita). Financial development is usually measured by the Credit-to-GDP ratio. Let’s suppose that we want to investigate the following:

$$(Credit/GDP)_{it} = \beta(GDP\ per\ capita)_{it} + u_{it} \tag{8}$$

A very extensive literature recognizes a ***positive*** relationship between the two variables, based broadly on microeconomic theories of credit market imperfections among other explanations (see for instance Djankok et al. (2007), Cottarelli et al. (2003) and Becerra et al. (2010)).

However, if the permanent income hypothesis holds, we can argue that if an increase in output (GDP) is perceived by economic agents to be only a short-term temporary change, consumers or investors have no reason to increase leverage in order to invest or consume more. Thus, a ***short-term*** increase in output should only affect the denominator in the Credit/GDP ratio and thus, it should have a ***negative*** impact on it.

In the same fashion, there could be a ***negative*** relation if households/firms increase credit levels to smooth consumption or investment at times when their income is temporarily below expected levels.

In summary, all these reasons imply that different components of the income per capita may have different effects on the credit ratio.

Let’s consider the following. We want to explain Y_{it} as in equation (1), but we have that:

$$X_{it} = f(Z_{it}^{1} \dots Z_{it}^{H}, Z_{it}^{H+1} \dots Z_{it}^{K}, Z_{it}^{K+1} \dots Z_{it}^{T}, e_{it}) \tag{9}$$

Where Z_{it}^{1} to Z_{it}^{H} are factors that are constant over time, Z_{it}^{H+1} to Z_{it}^{K} are factors that change slowly over time, and Z_{it}^{K+1} to Z_{it}^{T} are factors that changes widely over time and e_{it} is a noise component.

Then, we might decompose our explanatory variable X_{it} into two, three or more components (“Long-Term” or constant over time, Medium-Term and Cyclical or Short-Term):

$$X_{it} = X_{it}^{LT} + X_{it}^{MT} + X_{it}^{CYC} \tag{10}$$

And each of them could have its own different effect on the dependent variable:

$$Y_{it} = \beta^{LT} X_{it}^{LT} + \beta^{MT} X_{it}^{MT} + \beta^{CYC} X_{it}^{CYC} + u_{it} \tag{11}$$

Thus, equation (11) would be the *correct specification* that we should estimate empirically, provided that we can identify the underlying components. Consequently, if instead of (11) we were to estimate the following equation:

$$Y_{it} = \beta^{B} \bar{X}_{i\cdot} + \beta^{W}(X_{it} - \bar{X}_{i\cdot}) + u_{it} \tag{12}$$

We will have the following testable implications according to the previous basic theoretical assumptions:

A. If $\beta^{LT} = \beta^{MT} = \beta^{CYC}$ then $\beta^{B} = \beta^{W}$. However, if any of the former differs, then the latter should also differ. In other words if any of the components have a different effect, then the Between and the Within estimators should also be different.

B. The Between estimator β^{B} should be similar or equal to β^{LT}.

C. The Within estimator β^{W} should be *a combination* of the medium-term and short-term components' effects β^{MT} and β^{CYC} when *T* is large, and it should capture the short-term effect when *T* is small.

Proposition (C) follows from the fact that given our assumptions, $(X_{it} - \bar{X}_{i\cdot}) = X_{it}^{MT} + X_{it}^{CYC}$, so if we have a medium-term component that has a different effect than the short-term component, i.e. if $\beta^{MT} \neq \beta^{CYC}$ then the Within estimator would not be capturing a purely cyclical effect, but a combination of a medium-term effect and a cyclical one. However, this would only be true in a sample that covers a relatively long period of time, i.e. a long *T*, because in a sample with a short length, i.e. small *T*, the medium-term component could be undistinguishable from the long-term one.

Here it is also important to emphasize that $X_{it}^{LT} + X_{it}^{MT} + X_{it}^{CYC} = X_{it}$ and $\bar{X}_{i\cdot} + (X_{it} - \bar{X}_{i\cdot}) = X_{it}$ so if $\beta^{LT} = \beta^{MT} = \beta^{CYC}$ then estimating (11) or (12) would be exactly equivalent to estimating $Y_{it} = \beta X_{it} + u_{it}$. Therefore, the proposed approach should not have any consequence beyond increasing the number of regressors in the equation.

3 Empirical Exercise

In order to empirically test propositions A to C in the specific case of the relationship between private leverage and real income per capita, we estimate equations (8), (11) and (12) in a sample of 167 countries from 1990 to 2014, using different estimators, Random Effects, Within, Between, OLS and Arellano-Bond. The dependent variable is the private Credit-to-GDP ratio taken from the IFS (IMF) and the World Bank, and the explanatory variable is the GDP per capita measured in real US dollars and PPP terms. Its source is WEO (IMF)[7].

The specific hypotheses derived from our proposition that we want to test are the following:

A. The Between and the Within estimators should be statistically different. Moreover, if we decompose the GDP per capita into three components (long, medium and short term) their estimates should also be statistically different.

B. The Between estimate should be positive given that it is associated to the long-term component.

C. We also expect a positive sign for the effect of the medium-term component, although we do not have a prior on whether it could be larger or smaller than the long-term effect. In any case, the Within estimate should vary between a positive and a negative value depending on the length of the sample. In large T samples the Within estimate should be positive but smaller than the Between estimate, and in small T samples the Within estimate should be negative.

In all the regressions that we will run here the only explanatory variable will be the real GDP per capita or some variation of it, either its components or its lags. We refrain from including other explanatory variables in order to avoid diverting the discussion to whether we are omitting any specific variable. We have ran numerous robustness exercises including several other explanatory variables and the results related to the GDP per capita remain intact.

In Table 1 we show the results of estimating equation (8). We can also observe the Hausman test of the difference between the Within and Random Effects GLS estimates:

Table 1
Basic specification under different estimators

	Random GLS	Within	Between	OLS
Log GDP per capita	**29.575*** **	**31.396*** **	**23.496*** **	**24.338*** **
	(29.01)	(27.04)	(11.08)	(49.82)
Within R2	0.165	0.165	0.165	
Between R2	0.427	0.427	0.427	
Overall R2	0.391	0.391	0.391	0.391
Hausman Test p-value		0.001		

Source: BBVA Research

Given that in practice, besides the possible presence of individual fixed-effects we could also have time fixed-effects that bias our results, we also estimate equation (8) including time dummies for each year (Table 2).

7: The GDP per capita in PPP from WEO comes in current dollars. We use the USA GDP deflator to turn it into real dollars.

Table 2
Basic specification under different estimators including time dummies

	Random GLS	Within	Between	OLS
Log GDP per capita	**11.829*** **	**6.3351*** **	**23.496*** **	**23.767*** **
	(9.78)	(4.35)	(11.08)	(48.49)
Within R2	0.295	0.298	0.165	
Between R2	0.400	0.361	0.427	
Overall R2	0.366	0.270	0.391	0.403
Hausman Test p-value		0.0077		

Source: BBVA Research

The results clearly show that Between and the Within estimators are statistically different according to the standard Hausman test. This would imply that we should hold to the Within estimated effect according to the standard textbook suggestion. Additionally, once we control for time effects, the Within estimate is clearly *smaller* than the Between or the Random Effects estimates.

In the same fashion, we estimate equation (12) first without including time dummies (Table 3) and then including them (Table 4). In this specification the explanatory variables are:

- Log GDP pc (Country Mean): $\bar{X}_{i\cdot}$
- Log GDP pc (deviation from country mean): $(X_{it} - \bar{X}_{i\cdot})$

Table 3
Specification with two components of GDP per capita

	Random GLS	Within	Between	OLS
Log GDP per capita (Country Mean)	**23.452*** **		**23.381*** **	**23.945*** **
	(10.97)		(10.92)	(48.04)
Log GDP per capita (deviation from country mean)	**31.395*** **	**31.396*** **		**30.396*** **
	(27.05)	(27.04)		(12.06)
Within R2	0.165	0.165	0.163	
Between R2	0.427	0.011	0.424	
Overall R2	0.392	0.027	0.388	0.3895
Hausman Test p-value	0.883			
Test X1 = X2 (p-value)	0.000			

Source: BBVA Research

Table 4
Specification with two components of GDP per capita, including time dummies

	Random GLS	Within	Between	OLS
Log GDP per capita (Country Mean)	**23.595*** **		**23.734*** **	**24.095*** **
	(10.84)		(10.93)	(47.72)
Log GDP per capita (deviation from country mean)	**6.4946*** **	**6.3351*** **		**15.466*** **
	(4.46)	(4.35)		(4.92)
Within R2	0.298	0.298	0.163	
Between R2	0.410	0.000	0.424	
Overall R2	0.402	0.042	0.388	0.4045
Hausman Test p-value	0.311			
Test X1 = X2 (p-value)	0.000			

Source: BBVA Research

Not surprisingly, we obtain basically the same results than before. Moreover, we always reject the Hypothesis that $\beta^B = \beta^W$ when estimating equation (12) in the Random Effects and the OLS case.

In order to explicitly test for different effects of long, medium and short-term components, we further estimate a specification like equation (11) using a 5 years moving average of the GDP per capita as a measure of the medium term component and the deviation from the moving average as a measure of the cyclical component. Therefore, the explanatory variables are the following:

- Log GDP pc (Country Mean): Long term component: $\bar{X}_{i.}$
- Log GDP pc (MA 5 yrs - Country Mean): Medium term component: $(\bar{X}_{it}^{5yrs} - \bar{X}_{i.})$
- Log GDP pc (obs – MA 5 yrs): Short term component: $(X_{it} - \bar{X}_{it}^{5yrs})$

Table 5
Specification with three components of GDP per capita

	Random GLS	Within
Log GDP per capita (Country Mean)	**23.551*****	
	(11.53)	
Log GDP pc (MA 5 yrs - Country Mean)	**39.002*****	**39.137*****
	(27.67)	(27.79)
Log GDP pc (obs - MA 5 yrs)	**-24.15*****	**-25.10*****
	(-4.58)	(-4.76)
Within R2	0.205	0.205
Between R2	0.424	0.424
Overall R2	0.395	0.392
Hausman Test p-value	0.000	
Test X1 = X2 (p-value)	0.000	0.000

Source: BBVA Research

Table 6
Specification with three components of GDP per capita, including time dummies

	Random GLS	Within
Log GDP per capita (Country Mean)	**23.521*****	
	(10.71)	
Log GDP pc (MA 5 yrs - Country Mean)	**10.069*****	**9.9156*****
	(6.81)	(6.69)
Log GDP pc (obs - MA 5 yrs)	**-33.89*****	**-33.78*****
	(-8.07)	(-8.03)
Within R2	0.317	0.317
Between R2	0.413	0.004
Overall R2	0.405	0.050
Hausman Test p-value	0.000	
Test X1 = X2 (p-value)	0.000	0.000

Source: BBVA Research

In Table 5 we show the results without including time dummies and in Table 6 including them. In both cases the results are clearly as expected. The effect of the cyclical component is *negative*, as the permanent income hypothesis would suggest. On the other hand, the long-term and the medium-term components effects turn out to be *positive*, also in accordance to the expected theoretical implications.

However, when we do not include time dummies the effect of the medium-term component is higher than the long-term one, meanwhile, when we control for time dummies the opposite happens.

As a robustness exercise, we also estimate equation (11) but replacing the 5 years moving average with a Hodrick-Prescott estimated trend as a measure of the medium-term component and its deviation as a measure of the cyclical one:

- Log GDP pc (Country Mean): Long term component: $\bar{X}_{i.}$
- Log GDP pc (Trend HP): Medium term component: $(\bar{X}_{it}^{HP} - \bar{X}_{i.})$
- Log GDP pc (Cycle HP): Short term component: $(X_{it} - \bar{X}_{it}^{HP})$

Table 7
Specification with three components of GDP per capita (HP trend)

	Random GLS	Within	Random GLS
Log GDP per capita (Country Mean)			**40.640*** **
			(9.03)
Log GDP per capita (Trend HP)	**49.777*** **	**51.652*** **	**51.684*** **
	(26.72)	(25.23)	(25.25)
Log GDP per capita (Cycle HP)	**-0.509**	**-0.557**	**2.7850**
	(-0.09)	(-0.09)	(0.45)
Within R2	0.265	0.265	0.265
Between R2	0.447	0.447	0.440
Overall R2	0.418	0.418	0.416
Hausman Test p-value	0.261		
Test X1 = X2 (p-value)	0.000	0.000	0.026

Source: BBVA Research

Table 8
Specification with three components of GDP per capita (HP trend), including time dummies

	Random GLS	Within	Random GLS
Log GDP per capita (Country Mean)			**40.604*** **
			(8.9)
Log GDP per capita (Trend HP)	**19.922*** **	**9.0710*** **	**9.2375*** **
	(7.44)	(2.75)	(2.8)
Log GDP per capita (Cycle HP)	**-13.80** **	**-26.58*** **	**-25.66*** **
	(-2.32)	(-4)	(-4.08)
Within R2	0.351	0.354	0.354
Between R2	0.388	0.008	0.434
Overall R2	0.369	0.030	0.428
Hausman Test p-value	0.190		
Test X1 = X2 (p-value)	0.000	0.000	0.000

Source: BBVA Research

The results in Tables 7 and 8 show that the differences in the components effects are robust to an alternative definition of the medium-term and the cyclical component.

Importantly, the results obtained through the Within estimator that include time dummies (Table 8) are quite important because they show that we obtain *different effects* of the medium and short-term components while using an estimator that is *consistent* to the presence of both time and individual effects. The use of the Within estimator controls for the presence of individual fixed effects and the additional use of time dummies controls for the presence of time fixed effects.

Finally we also estimate equation (11) specification using the Arellano-Bond estimator. In this case we are controlling for the presence of individual and time fixed effects and we are additionally using a dynamic specification.

Table 9
Specification with medium (MA 5 years) and short-term components using Arellano-Bond estimator

	Without Time Dummies	Including Time Dummies	Implied Long Run	Implied Long Run (Time Dummies)
Y (t-1)	**0.8572*** **	**0.8214*** **		
	(74.87)	(62.99)		
Log GDP pc (MA 5 yrs - Country Mean)	**10.474*** **	**4.4821*** **	**73.38**	**25.10**
	(11.27)	(3.94)		
Log GDP pc (obs - MA 5 yrs)	**12.764*** **	**-4.489* **	**89.42**	**-25.14**
	(5.52)	(-1.84)		
Overall R2	0.946	0.981		
Test X1 = X2 (p-value)	0.000	0.000		

Source: BBVA Research

The results in Table 9 show that when we estimate a dynamic specification we can also obtain different effects of the medium and short-term components (positive and negative). Since the Arellano-Bond with time dummies is a consistent estimator and explicitly accounts for a dynamic specification, it shows that the estimated difference in the coefficients cannot be explained by the existing theoretical explanations.

As an additional robustness check we repeat all the same estimations performed in this section but using other non-lineal specifications, trying to account for the possible existence of different non-linear functional forms in the relationship between the credit ratio and the income per capita (Poisson and Gompertz Curve). Although we do not show the results, we can simply say that the results are identical in terms of the estimated signs and the relative size of the coefficients.

4 Monte Carlo Simulation Exercise

In order to evaluate whether the theoretical explanation based on the misspecification of the dynamics could explain the previous empirical results we run a simple Monte Carlo simulation in which we simulate different samples of artificial data. We first generate 50 samples of 5000 observations with 200 individuals (N=200) and 25 years (T=25). We also generate 50 samples of 500 observations (N=20 and T=25) in order to test the effect of a small number of individuals (small N). Finally, we generate 50 samples of 800 observations (N=200 and T=4) in order to identify the effects of having a short time-series (small T)[8].

The explanatory variables are simulated according to our theoretical assumptions. First, we generate 3 *explanatory* variables X1, X2 and X3 that include three types of underlying factors: i) Two factors that are constant over time but that differ across individuals; ii) Two factors that move slowly over time because they are simulated as a moving average of other normal random variables; and iii) two factors that are just random normal variables in order to reproduce variables that vary widely over time.

After generating the explanatory variables, we simulate the long-term, the medium-term and the cyclical component through a Hodrick-Prescott filter on each one of the explanatory variables (X1, X2 and X3). With these components, we then generate our dependent variables. As a robustness exercise, we also decompose the variables using a moving average of 5 years.

For each set of samples we generate two dependent variables that come from two different specifications. First, one specification in which the explanatory variables are allowed to have three different components with their own specific effects, i.e. equation (11). Secondly, a specification in which we include multiple lags of the explanatory variables as in equation (5).

The true parameters for each one of the explanatory variables in each specification are shown in Tables 10 and 11[9]:

Table 10
True values of the coefficients of the three of each explanatory variable

	β (LT)	β (MT)	β (CYC)
X1	2.7	1.7	-0.5
X2	3.5	1.5	0.85
X3	1.5	1.5	3.5

Source: BBVA Research

Table 11
True values of the coefficients of components each lag of the explanatory variables

	β (t)	β (t-1)	β (t-2)	β (t-3)	β (t-4)	β (t-5)	Long Run
X1	1.5	1.3	1.1	0.8	0.7	0.6	6
X2	1.8	1.6	1.3	1	0.8	0.7	7.2
X3	2.5	2.5					5

Source: BBVA Research

For instance, for the case of variable X1, the components specification would be:

$$Y_{it} = 2.7\bar{X}_{i\cdot} + 1.7(\bar{X}_{it}^{HP} - \bar{X}_{it}) - 0.5\left(X_{it} - \bar{X}_{it}^{HP}\right) + \cdots + u_{it}$$

And the dynamic specification would be:

$$Y_{it} = 1.5X_{it} + 1.3X_{it-1} + 1.1X_{it-2} + 0.8X_{it-3} + 0.7X_{it-4} + 0.6X_{it-5} + \cdots + u_{it}$$

We clarify that the specification in each case the three variables X1, X2 and X3 are included.

8: We also run the same simulation for T=5 and T=3, with the same results.
9: We have defined the true parameters of the lags so that the long-term response could be roughly approximated by the Arellano-Bond estimator using the lag dependent variable coefficient ρ. The lag-coefficients are thus defined as $\beta_r = \rho^r \beta$ with ρ=0.75, and β=1.75, where β is the short-term (contemporaneous) response and *r* corresponds to the lag period.

Through this simulation exercise we will try to identify which one of the two misspecifications produces estimates that are more consistent with the empirical differences found in the Between and the Within estimators, and the different signs of the short-term and long-term components.

Therefore, we estimate the Random Effects, the Within estimator and the Arellano Bond estimator in all the simulated samples, but purposely ignoring the true specifications in both cases, i.e. including only the explanatory variables X1, X2 and X3 instead of the true explanatory variables: the three components in the first case and the different lags in the second one. We will denote the first misspecification the “components misspecification” and the second one “dynamic misspecification”.

We should clarify that in the case of multiple lags (dynamic misspecification) the Arellano-Bond estimator is not exactly miss-specified since it allows us to consistently estimate the short-run response.

The results of the simulated data regressions for a long sample (T= 25) are shown in Table 12 and Table 13. The results for a short sample period (T=4) are shown in Tables 14 and 15.

Table 12
Long sample (T=25) component misspecification

	Between	Within	Random	Arellano-Bond	AB Implied Long Run
X1	**2.7*** **	**1.1*** **	**1.94*** **	**0.13*** **	**0.2**
	(360.6)	(69.3)	(106.1)	(4.8)	
X2	**3.5*** **	**1.29*** **	**2.6*** **	**1.07*** **	**1.6**
	(303)	(46.5)	(87)	(24.6)	
X3	**1.5*** **	**2.01*** **	**1.66*** **	**3.04*** **	**4.4**
	(134.9)	(62.8)	(54)	(53.1)	
Z1	**0.34** **	**0.3*** **	**0.29*** **	**0.32*** **	
	(2.4)	(5.1)	(3)	(6.4)	
Z2	**0.41*** **	**0.4*** **	**0.4*** **	**0.43*** **	
	(2.8)	(6.8)	(4.1)	(8.6)	
Y (t-1)				**0.31*** **	
				(20.8)	
Within R2	0.57	0.70	0.62		
Between R2	0.999	0.77	0.98		
Overall R2	0.80	0.71	0.81	0.57	
Hausman Test p-value	0.000				

Source: BBVA Research

Table 13
Long sample (T=25) and dynamic misspecification

	Between	Within	Random	Arellano-Bond	AB Implied Long Run
X1	**6.09*** **	**4.46*** **	**5.01*** **	**1.59*** **	**5.2**
	(68.5)	(72.5)	(94.1)	(37.4)	
X2	**7.27*** **	**4.88*** **	**5.8*** **	**1.8*** **	**5.8**
	(53.9)	(46.8)	(66.9)	(27.3)	
X3	**5*** **	**4.7*** **	**4.85*** **	**2.44*** **	**7.9**
	(38.6)	(37.3)	(51.1)	(24.4)	
Z1	**0.05**	**0.28**	**0.29**	**0.4*** **	
	(0)	(1.6)	(1.6)	(5.9)	
Z2	**0.39**	**0.41** **	**0.4** **	**0.52*** **	
	(0.3)	(2.3)	(2.2)	(7.7)	
Y (t-1)				**0.69*** **	
				(108.9)	
Within R2	0.69	0.70	0.70		
Between R2	0.979	0.96	0.97		
Overall R2	0.88	0.88	0.89	0.99	
Hausman Test p-value	0.000				

Source: BBVA Research

Table 14
Short sample (T=4) and components misspecification

	Between	Within	Random	Arellano-Bond	AB Implied Long Run
X1	**2.32*** **	**-0.07**	**1.05** **	**-0.16** **	**-0.16**
	(33.1)	(-1.1)	(-2.6)	(-2.6)	
X2	**3.02*** **	**0.94*** **	**1.86*** **	**0.94*** **	**0.96**
	(27.5)	(9.6)	(10.2)	(10.2)	
X3	**1.52*** **	**3.04*** **	**2.1*** **	**3.26*** **	**3.35**
	(14)	(21.1)	(24.3)	(24.3)	
Z1	**0.24**	**0.28*** **	**0.25*** **	**0.28*** **	
	(0.4)	(2.9)	(3.1)	(3.1)	
Z2	**0.49**	**0.42*** **	**0.41*** **	**0.41*** **	
	(0.8)	(4.2)	(4.6)	(4.6)	
Y (t-1)				**0.03**	
				(0.6)	
Within R2	0.09	0.34	0.26		
Between R2	0.91	0.19	0.75		
Overall R2	0.83	0.20	0.71	0.19	
Hausman Test p-value	0.000				

Source: BBVA Research

Table 15
Short sample (T=4) and dynamic misspecification

	Between	Within	Random	Arellano-Bond	AB Implied Long Run
X1	**5.83*** **	**1.82*** **	**3.94*** **	**1.45*** **	**4.97**
	(40.7)	(12.2)	(14.1)	(14.1)	
X2	**7.03*** **	**1.75*** **	**4.4*** **	**1.68*** **	**5.74**
	(31.2)	(7.9)	(11)	(11)	
X3	**4.99*** **	**3.27*** **	**4.45*** **	**2.46*** **	**8.40**
	(22.4)	(9.9)	(10.3)	(10.3)	
Z1	**0.05**	**0.34**	**0.34*** **	**0.43*** **	
	(0)	(1.6)	(2.9)	(2.9)	
Z2	**0.44**	**0.43***	**0.44*** **	**0.52*** **	
	(0.4)	(1.9)	(3.5)	(3.5)	
Y (t-1)				**0.71*** **	
				(22.9)	
Within R2	0.30	0.35			
Between R2	0.94	0.81			
Overall R2	0.89	0.78	0.88	0.98	
Hausman Test p-value	0.000				

Source: BBVA Research

Although we do not show the results for the small N simulation, it suffices to say that the number of individuals (N) does not have any effect in our simulated estimates.

We can summarize the simulation predictions shown in Tables 14 to 17 in the following four parts:

1. *Long series (large T) and components misspecification*:
   a. The Between estimator correctly estimates the true long-term component effect.
   b. The Within estimator yields a combination of the true medium and short-term components' effects (it is close to the sum of the two).
   c. The Arellano-Bond estimator yields also a combination of the medium-term and short-term components, but is much closer to the true short-term effect than the Within estimator. In any case we cannot recover the true long-term effect from the Arellano-Bond estimates.
2. *Long series (large T) and dynamic misspecification*:
   a. The Between effect captures exactly the true long-term responses.
   b. The Within estimator is severely upward bias with respect to the true short-run response and it is actually closer to the long-term response, something that is in line with the results of Eggen et al (2005).
   c. Finally, the Arellano-Bond estimator correctly estimates the true short-run response (the contemporary effect) and it approximates well the true long-run response.
3. *Short series (small T) and components misspecification*:
   a. The Between estimate becomes slightly downward biased with respect to the true long-term component effect, but only in the X1 and X2 cases, when the true long-term was much higher than the true medium-term. In the case of X3 is unbiased.

b. The Within estimator now becomes much closer to the true short-term component effect. For instance, in the case of variable X1 that has by construction a negative short-term effect, the Within estimate correctly becomes negative as well, although is upward biased.

c. The Arellano-Bond estimator yields also a combination of the true medium-term and short-term components, but in this case comes very close to the true *short-term effect*. Again, we cannot recover the true long-term effect or the medium-term effect from the Arellano-Bond estimates.

4. *Short series (small T) and dynamic misspecification*:

   a. The Between estimator is unbiased with respect to the true long run response.

   b. The Within estimator still shows an upward bias with respect to the short-run response but now is closer to the short response than the long response compared to the large T case. Again this is in line with the results of Eggen et al (2005).

   c. The Arellano-Bond estimator again correctly estimates the short-run response and importantly, the estimates are exactly the same as in the large T case. This is important because it show us that the Arellano-Bond estimates do not vary with the length of the sample when we are estimating a relationship that comes from a dynamic specification.

Our simulation results together with those from Eggen and Pfaffermayr (2005) provide clear support for our theoretical explanation of the empirical findings versus the alternative explanation based on a dynamic misspecification.

First of all, our simulations and those from Eggen and Pfaffermayr (2005) indicate that in the dynamic misspecification case, given a positive true short run parameter, the Within estimate is *upward* biased in both the small T and the large T cases (if the memory of the explanatory variable is high, as in the case of GDP per capita). Therefore, it follows that given that in the small T case we find a *negative* Within estimate, this means that in reality the true long-term effect should be *lower* because of the upward bias. Thus the true long-run response would have to be *more negative* than the short-run response, something that conflicts with the *positive* Between estimate that we obtain in the empirical exercise.

Also importantly, although the simulated Within coefficients decrease when we reduce the length of the time series (T) in both type of misspecifications, the simulated Arellano-Bond is not sensitive to the length of the time series in dynamics misspecification case. However, in the components misspecification simulation, the Arellano-Bond is indeed sensible to the number of periods. More specifically, when we go from a long sample to a short sample, the Arellano-Bond estimate goes from a combination of the medium-term and the short-term, to being more closely associated to the short-term component. This means that in reality, if the true medium-term effect is positive and the true short-term effect is negative, the Arellano-Bond estimator would change from a positive to a negative estimate.

In order to test these simulation predictions with the data, we estimate the Between, the Within and the Arellano-Bond coefficients with our empirical data in smaller subsamples. Since our data sample consists of 25 years, we estimate the coefficients in several subsamples of 5, 4 and 3 years long, and we compare the average of all the subsamples to the estimated coefficients obtained using the 25 years (whole sample)[10].

In each regression we estimate equation (8), i.e. the credit-ratio as a function of the log of the GDP per capita. These regressions only include the GDP per capita, nor the components, nor any lag. In tables 16 and 17 we can observe the average coefficients estimated in all these subsamples.

10: By splitting the 25 years from the original sample (1990 to 2014), we are able to estimate the parameters in 15 subsamples of 5 years using different combinations of years (1990-1994, 1991-1995, 1992-1996, etc.), 18 subsamples of 4 years, and 24 subsamples of 3 years.

We can see that in all the cases, both the Within and the Arellano-Bond estimates clearly *decrease* as we reduce the length of the samples (as T decreases). Moreover, in the Arellano-Bond specification that includes time dummies the estimated coefficient of the GDP per capita goes from a positive to a negative value.

This effect can also be observed if we estimate the proportion of subsamples in which the estimated effect is negative and we can see that it clearly grows as we reduce the sample period and as we go from the Within estimator to the Arellano-Bond estimator with time dummies (the theoretically more consistent and well-defined one).

Notably, the Between estimate does not change at all when we reduce the length of the time series. The simulation results predicted that it could decrease slightly, depending on the true difference between the long-term and the medium-term.

Table 16
Average estimated coefficients of the log of GDP per capita in different subsamples

	Whole Sample	Average Sub-samples 5 years	Average Sub-samples 4 years	Average Sub-samples 3 years
Within Estimator	31.4	17.0	12.2	9.8
Arellano Bond Estimator	12.4	6.2	5.4	4.0
AB Implied Long Run	79.5	14.2	10.9	9.1
Within + Time dummies	6.3	5.1	3.0	1.5
Arellano Bond + Time dummies	4.5	-1.5	-1.9	-2.7
AB Implied Long Run (+ T. Dum)	19.7	-3.6	-4.2	-5.9
Between Estimator	23.9	23.3	23.6	23.7

Source: BBVA Research

Table 17
Proportion of subsamples with negative coefficients of the GDP per capita

	Whole Sample	Average Sub-samples 5 years	Average Sub-samples 4 years	Average Sub-samples 3 years
Within Estimator	0%	8%	0%	0%
Arellano Bond Estimator	13%	25%	29%	13%
AB Implied Long Run	33%	42%	38%	33%
Within + Time dummies	20%	17%	46%	20%
Arellano Bond + Time dummies	53%	67%	71%	53%
AB Implied Long Run (+ T. Dum)	64%	72%	75%	64%
Between Estimator	0%	0%	0%	0%

Source: BBVA Research

5 Conclusions

We present a rationale and supporting evidence that suggests that the observed difference in the Between and the Within estimators could be due to actual different effects of different underlying components of an explanatory variable. More specifically, we present evidence that the structural (long-term) and short-term components of a variable can have different effects on the dependent variable we want to explain.

We also show through the use of Monte Carlo simulations that this explanation fits better the empirical results obtained when we study the relationship between private Credit-to-GDP and real income per capita than the alternative explanation based on the misspecification of the dynamics.

Thus, the Between estimator could indeed be capturing the effect of a long-term component, whereas the Within estimator could be capturing the effect of medium-term and short-term components, depending on the length of the sample (T).

The existence of such possible different effects of these components is quite important because it could help us reconcile many conflicting empirical results. For instance, when the long-term component has a significant effect but the others have a non-significant effect, using only the Within estimator might result in concluding that the variable is not significant at all.

It can also help us to improve the goodness of fit of our estimations, and we could also estimate and forecast the structural and cyclical levels of the dependent variable, when the distinction between the structural and cyclical parts of a variable and their respective determinants is of interest.

Moreover, if the theoretical situation presented here were actually a common phenomenon in empirical applications, it may force us to re-evaluate the meaning and the use of some specification tests widely used in the panel data analysis, such as the Hausman test. Simply, in the cases that the Between and the Within estimators were indeed capturing the effect of different components, then the tests that intend to identify whether the “Random Effects” or the Between estimator are not different from the Within (consistent) estimator would be doomed to fail.

Furthermore, both the Within (Fixed Effects) and the Arellano-Bond estimators are probably the most widely used in Panel Data applications. However, a widely discussed problem in the field is that these estimators wipe out the “between” variance of the variables, which severely complicates the estimation of the effect of time-invariant variables or those that have very little time variation.

Moreover, in several applied cases, the “between” variance represents the bulk of the total variance and when using the Within or other estimators that apply the same transformation we could be ignoring the lion’s share of the information contained in many variables. For instance, in the empirical data that we use here, the “between” variance represents 85% the total variance of the variable Credit-to-GDP ratio in our sample covering 25 years, from 1990 to 2014. If we estimate the same proportion in subsamples of 5 years, the average proportion rises up to 97%. In the case of the real GDP per capita the “between” variance represents 96% and 99.7% in the 25 years and 5 year subsample (in US$ constant PPP dollars).

Table 18
Percentage of the total variance explained by the Between variance

	Between Variance/Total Variance	
	Sample 25 years	**Sample 5 years**
Credit/GDP	85.53%	96.62%
GDP pc PPP in Logs	96.07%	99.71%

Source: BBVA Research

Bibliography

Balestra, P., Nerlove, M., 1966, Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas, *Econometrica* Vol. 34, No. 3, July 1966.

Baltagi, B.H. and J.M. Griffin, 1984, Short and long run effects in pooled models, *International Economic Review* 25, 631–645.

Baltagi, B.H., 2005, Econometric Analysis of Panel Data, John Wiley, Chichester, 4th edition.

Barajas,A., Beck, T., Dabla-Norris, E. and Yousefi, SR. (2013) "Too Cold, Too Hot, Or Just Right? Assessing Financial Sector Development Across the Globe", IMF Working Paper WP/13/81

Becerra, O., Cavallo, E. , Scartascini, C., 2010, The Politics of Financial Development: The Role of Interest Groups and Government Capabilities. APSA 2010 Annual Meeting Paper.

Claessens, S. and Laeven, L. (2004), "Financial Development, Property Rights, and Growth", Policy Research Working Paper World Bank 2924.

Cottarelli, C., Dell'Ariccia,G. and Vladkova-Hollar, I.(2003), "Early Birds, Late Risers, and Sleeping Beauties: Bank Credit Growth to the Private Sector in Central and Eastern Europe and the Balkans", IMF Working Paper WP/03/213.

Djankov, S., McLiesh, C. and Shleifer, A. (2007), "Private Credit in 129 Countries", Journal of Financial Economics 84 (2007) 299–329.

Doel van den, I.T., Kiviet, J.F., 1994, Asymptotic Consequences of Neglected Dynamics in Individual Effects Models, Statistica Neerlandica 48, pp. 71-85.

Doel van den, I.T., Kiviet, J.F., 1995, Neglected Dynamics in Panel Data Models; Consequences and Detection in Finite Samples, Statistica Neerlandica 49,pp. 343-361.

Egger, P. & Pfaffermayr, M. , 2005, Estimating Long Run and Short Run Effects in Static Panel Models, *Econometric Reviews* Volume 23, Issue 3, 2005

Griliches, Z. and J.A. Hausman, 1986, Errors in variables in panel data, *Journal of Econometrics* 31, 93–118.

Houthakker, H.S, 1965, New Evidence on Demand Elasticities, *Econometrica* 33, pp. 277-288.

Kuh, E., 1959, The Validity of Cross-sectionally Estimated Behavior Equations in Time Series Applications, *Econometrica* 27, 197-214

Mairesse, J., 1990, Time-series and cross-sectional estimates on panel data: Why are they different and why should they be equal?, in J. Hartog, G. Ridder and J. Theeuwes, eds., *Panel Data and Labor Market Studies* (North-Holland, Amsterdam), 81–95.

Mairesse, J., Sassenou, M., 1991, R&D Productivity: A Survey of Econometrics Studies at the Firm Level, NBER Working Paper No. 3666.

Mundlak, Y., 1978, On the pooling of time series and cross-section data, *Econometrica* 46, 69–85.

Pirotte, A., 1999, Convergence of the static estimation toward long run effects of dynamic panel data models, *Economics Letters* 53, 151–158.

Stata 12 Longitudinal Data/Panel Data Reference Manual, Release 12. College Station, TX: StataCorp LP.

Working Papers

2016

16/10 **Alfonso Ugarte Ruiz:** Long run and short run components in explanatory variables and differences in Panel Data estimators.

16/09 **Carlos Casanova, Alicia García-Herrero:** Africa's rising commodity export dependency on China.

16/08 **Ángel de la Fuente:** Las finanzas autonómicas en 2015 y entre 2003 y 2015.

16/07 **Ángel de la Fuente:** Series largas de algunos agregados demográficos regionales, 1950-2015.

16/06 **Ángel de la Fuente:** Series enlazadas de Contabilidad Regional para España, 1980-2014.

16/05 **Rafael Doménech, Juan Ramón García, Camilo Ulloa:** Los efectos de la flexibilidad salarial sobre el crecimiento y el empleo.

16/04 **Angel de la Fuente, Michael Thöne, Christian Kastrop:** Regional Financing in Germany and Spain: Comparative Reform Perspectives.

16/03 **Antonio Cortina, Santiago Fernández de Lis:** El modelo de negocio de los bancos españoles en América Latina.

16/02 **Javier Andrés, Ángel de la Fuente, Rafael Doménech:** Notas para una política fiscal en la salida de la crisis.

16/01 **Ángel de la Fuente:** Series enlazadas de PIB y otros agregados de Contabilidad Nacional para España, 1955-2014.

2015

15/33 **Shushanik Papanyan:** Digitization and Productivity: Where is the Growth? Measuring Cycles of Technological Progress.

15/32 **Alfonso Arellano, Noelia Cámara, David Tuesta:** Explaining the Gender Gap in Financial Literacy: the Role of Non-Cognitive Skills.

15/31 **Ángel de la Fuente:** Series enlazadas de Contabilidad Regional para España, 1980-2014. Parte II: Empleo asalariado, rentas del trabajo y salarios medios.

15/30 **Jingnan Cai, Alicia García-Herrero, Le Xia:** Regulatory arbitrage and window-dressing in the shadow banking activities: evidence from China's wealth management products.

15/29 **Javier Alonso, Alfonso Arellano:** Heterogeneity and diffusion in the digital economy: spain's case.

15/28 **Javier Alonso, Alfonso Arellano:** Heterogeneidad y difusión de la economía digital: el caso español.

15/27 **Angel de la Fuente**: Series enlazadas de Contabilidad Regional para España, 1980-2014.

15/26 **Carlos Casanova, le Xia and Romina Ferreira:** Measuring Latin America's export dependency on China.

15/25 **Nathaniel Karp, Boyd W. Nash-Stacey:** Embracing the Financially Excluded in the U.S.: A Multi-Dimensional Approach to Identifying Financial Inclusion Across MSAs and Between Cohorts.

15/24 **Alicia Garcia-Herrero, K.C. Fung:** Determinants of Trade in Parts and Components: An Empirical Analysis.

15/23 **Mariano Bosch, Angel Melguizo, Enith Ximena Peña, David Tuesta:** El ahorro en condiciones formales e informales.

15/22 **Antonio Villar:** Crisis, households' expenditure and family structure: The Palma ratio of the Spanish economy (2007-2014).

15/21 **Andrés Hernández, Bernardo Magnani, Cecilia Posadas, Jorge Redondo, Gonzalo Robles, Juan M. Ruiz y Enestor Dos Santos:** ¿Cuáles son los sectores con mayor potencial para aprovechar la Alianza del Pacífico?

15/20 **Gonzalo de Cadenas, Alicia Garcia-Herrero, Alvaro Ortiz and Tomasa Rodrigo:** An Empirical Assessment of Social Unrest Dynamics and State Response in Eurasian Countries. / *Published in Eurasian Journal of Social Sciences, 3(3), 2015, 1-29.*

15/19 **Mariano Bosch, Angel Melguizo, Enith Ximena Peña and David Tuesta:** Savings under formal and informal conditions.

15/18 **Alicia Garcia-Herrero, K.C. Fung, Jesus Seade:** Beyond Minerals: China-Latin American Trans-Pacific Supply Chain.

15/17 **Alicia Garcia-Herrero, Le Xia, Carlos Casanova:** Chinese outbound foreign direct investment: How much goes where after round-tripping and offshoring?

15/16 **Diego José Torres Torres:** Evaluando la capacidad predictiva del MIDAS para la Eurozona, Alemania, Francia, Italia y Portugal.

15/15 **Alicia Garcia-Herrero, Eric Girardin, Arnoldo Lopez-Marmolejo:** Mexico's monetary policy communication and money markets.

15/14 **Saidé Salazar, Carlos Serrano, Alma Martínez, Arnulfo Rodríguez**: Evaluation of the effects of the Free Trade Agreement between the European Union and Mexico (EU-MX FTA) on bilateral trade and investment.

15/13 **Saidé Salazar, Carlos Serrano, Alma Martínez, Arnulfo Rodríguez:** Evaluación de los efectos del Tratado de Libre Comercio entre la Unión Europea y México (TLCUEM) en el comercio bilateral y la inversión.

15/12 **Alicia Garcia-Herrero, Eric Girardin and Enestor Dos Santos:** Follow what I do, and also what I say: Monetary policy impact on Brazil's financial markets.

15/11 **Noelia Cámara, David Tuesta, Pablo Urbiola:** Extendiendo el acceso al sistema financiero formal: el modelo de negocio de los corresponsales bancarios.

15/10 **Noelia Cámara, David Tuesta, Pablo Urbiola:** Extending access to the formal financial system: the banking correspondent business model.

15/09 **Santiago Fernández de Lis, José Félix Izquierdo de la Cruz y Ana Rubio González:** Determinantes del tipo de interés del crédito a empresas en la Eurozona.

15/08 **Pau Rabanal and Juan F. Rubio-Ramírez:** Can International Macroeconomic Models Explain Low-Frequency Movements of Real Exchange Rates?.

15/07 **Ándel de la Fuente y Rafael Doménech:** El nivel educativo de la población en España y sus regiones: 1960-2011.

15/06 **Máximo Camacho and Jaime Martínez-Martín:** Monitoring the world business cycle. / *Published in Economic Modelling 51 (2015) 617–625.*

15/05 **Alicia García-Herrero and David Martínez Turégano:** Financial inclusion, rather than size, is the key to tackling income inequality.

15/04 **David Tuesta, Gloria Sorensen, Adriana Haring y Noelia Cámara:** Inclusión financiera y sus determinantes: el caso argentino.

15/03 **David Tuesta, Gloria Sorensen, Adriana Haring y Noelia Cámara:** Financial inclusion and its determinants: the case of Argentina.

15/02 **Álvaro Ortiz Vidal-Abarca and Alfonso Ugarte Ruiz:** Introducing a New Early Warning System Indicator (EWSI) of banking crises.

15/01 **Alfonso Ugarte Ruiz:** Understanding the dichotomy of financial development: credit deepening versus credit excess.

Click here to Access the Working Paper published

Spanish

and English

Contact details:
BBVA Research
Azul Street, 4
La Vela Building - 4 and 5 floor
28050 Madrid (Spain)
Tel.: +34 91 374 60 00 and +34 91 537 70 00
Fax: +34 91 374 30 25
bbvaresearch@bbva.com
www.bbvaresearch.com