Dollarization Persistence and Individual Heterogeneity *

Paul Castillo
London School of Economics &
Central Bank of Peru
Jr. Miroquesada 441, Lima 1, Peru
pcastillob@bcrp.gob.pe

Diego Winkelried†
St John's College,
University of Cambridge
CB2 1TP, Cambridge, UK
dw295@cam.ac.uk

December 15, 2005

Abstract

The most salient feature of financial dollarization, and the one that causes more concern to policymakers, is its persistence: even after successful macroeconomic stabilizations, dollarization ratios often remain high. In this paper we claim that this persistence is connected to the fact that the participants in the dollar deposit market are fairly heterogenous, and so is the way they form their optimal currency portfolios. We develop a simple model when agents differ in their ability to process information, which turns out to be enough to generate persistence upon aggregation. We find empirical support for this claim with data from three Latin American countries and Poland.

Keywords : Dollarization, individual heterogeneity, persistence, aggregation.
JEL Codes : C43, E50, F30.

*The views expressed in this paper are those of the authors. Diego Winkelried gratefully acknowledges financial assistance from the ORS Scheme and the Cambridge Gates Trust.

†Corresponding author: ++ 44 7765 574 055.

1 Motivation

Even though dollarization is a relatively new research area, the experiences of many Latin American and transition economies during the 1990's has inspired a growing and rich body of related literature.[1] Dollarization is normally associated with the partial substitution of the domestic currency by a foreign currency (the US dollar) as a store of value, as opposed to currency substitution which refers to the use of the foreign currency as a medium of exchange.

In this paper by dollarization we mean *deposit dollarization*[2] which leads ultimately to credit dollarization and hence to the vulnerability of the financial system of highly dollarized countries. As stressed by Cook (2004) and Céspedes *et. al* (2004), the efficacy of monetary policy in small open economies with flexible exchange rates is compromised by the negative balance sheet effects generated by dollarization. In this case, sudden real depreciations can have detrimental effects on the economic activity by reducing the net worth of firms and generating adverse effects on investment. This situation gives a rationale for a "fear of floating" behavior of central banks (Calvo and Reinhart, 2002; Morón and Winkelried, 2005).

One of the most salient features of dollarization, and probably the one that causes more concern to policymakers, is its *persistence*. It is well documented that dollarization increases sharply during episodes of unduly macroeconomic instability and that it remains stubbornly high even after successful stabilizations.[3] A top-of-mind explanation of the hysteresis is lack of confidence in domestic currency assets as a result of the traumas brought by past inflation, devaluations, banking crises, and so on. This, however, is not very consistent with the strong macroeconomic fundamentals observed in several highly dollarized countries (notably Peru and some transition economies in the early 2000's).

An alternative approach to address this puzzle is to modify existing currency substitution models based on adjustment costs or network externalities. Guidotti and Rodríguez (1992), Sturzenegger (1997) and Uribe (1997) develop models in which the cost of using the dollar for transactions depends negatively on the currency substitution ratio, so once transactions get dollarized, there is no benefit to switch back to using domestic currency if others continue using dollars. An obvious limitation of this approach is that it refers to the medium-of-exchange and not to the store-of-value function of money. Furthermore, these models rely heavily on a "ratchet variable", a knowledge stock that drives the persistence, so although they explain neatly upward trends of the dollarization ratio, they do not seem useful in advising policymakers how to dedollarize, as this may imply an implausible reduction of the knowledge stock.

Ize and Levy Yeyati (2003) provide a different framework for modelling dollarization. They use

[1] See De Nicoló *et. al* (2005), Levy Yeyati (2006) and the references therein.

[2] Sometimes this is known as asset substitution (Reinhart *et. al*, 2003) or financial dollarization (Ize and Levy Yeyati, 2003).

[3] See Guidotti and Rodríguez (1992), Savastano (1996), Quispe (2000) and Kamin and Ericsson (2003).

a portfolio selection approach and derive a minimum variance portfolio (MVP) that depends on the relative volatility of inflation and real exchange depreciation rates. Dollarization would persist even when inflation is low and stable insofar as the volatility of real exchange depreciation is smaller than that of inflation. However, this framework is static whereas persistence is inherently a dynamic phenomenon. In our view, the MVP approach, which is by now very popular and has proven successful in explaining cross-sectional variation of dollarization,[4] was not designed to deal with dynamics, since the MVP, the underlying equilibrium level of dollarization, depends on *unconditional* moments.[5]

Curiously, a fact that researchers have apparently overlooked is the very nature of the participants of the dollar deposit market in dollarized economies: dollar depositors are extremely heterogenous, ranging from large entrepreneurs to small firms to non-profit organizations and to individuals (rich and not-so-wealthy).[6] Participation costs are virtually nil because of liberalization, deregulation and, importantly, due to the emergence of informal currency traders – known as *cambistas* in many Latin American countries – which benefit from buying and selling dollars with tighter markups than those in the banking sector.[7] A typical *cambista* would hold a limited amount of money for business (say, between US$3,000 and US$5,000) as she is aimed to meet the dollar demand for individuals or small firms, normally unwilling to pay the higher bank premium to get their savings dollarized.[8] All in all, participation becomes independent of the scale of the transaction and hence widespread.

The aim of this paper is to draw the attention to the fact that heterogeneity of depositors can easily explain the persistence of financial dollarization. As pointed out by Granger (1980), differences in individual dynamics lead to aggregate persistence. Thus, as it is reasonable to expect that the dynamics of the optimal currency portfolio of a financial expert differs from that of a blacksmith, a persistent aggregate dollarization ratio arises naturally. There are of course various differences between a financial expert and a blacksmith, but provided that both access the dollar deposit market almost for free, the relevant difference to our analysis centers in their ability to process information and, therefore, to make informed saving decisions.[9]

We provide a simple extension to the MPV approach by considering that depositors are

[4] Ize and Levy Yeyati (2003) provide empirical evidence that the MVP has some explanatory power for the average level of dollarization across countries. De Nicoló *et. al* (2005) extends this empirical analysis by considering a broader set of countries.

[5] Dollarization hysteresis is observed in several countries with high real exchange rate volatility, e.g. Russia. The reason of this apparent contradiction with the portfolio approach may be that it is very difficult to get a sound estimate of the unconditional variances that compose the MVP.

[6] An exception is Sturzenegger (1997) who studies the implications of income inequality on currency substitution, yet with no reference to deposit dollarization.

[7] Agénor and Haque (1996) provide an account of informal currency markets.

[8] Even large firms may find it profitable to trade with a pool of (well-organized) *cambistas*.

[9] Surely, income differences can also be important if the income gap between the financial expert and the blacksmith is wide. However, we find that in dollarized economies the dollar deposit participation of (many) firms and (a lot of) individuals can be taken roughly as having the same importance.

heterogeneous in their ability to process information and to forecast real returns. In particular, we assume that depositors receive noisy information and that the amount of noise they receive is idiosyncratic. It turns out that under these assumptions the individual's optimal dollarization ratio follows an AR(1) process, where its degree of persistence is associated to the ability of the depositor to extract information. In particular, the individual dollarization ratio will be less (more) persistent and it will respond more (less), on impact, to changes in the signal, higher (lower) the ability of the depositor to process information. Aggregating the individual dynamics of depositors we show that financial dollarization can be very persistent, even though at the individual level persistence is moderate, when heterogeneity across depositors is large

The plan of the paper is as follows. In section 2 we briefly explore these issues using Peruvian and Polish data.[10] For reasons explained below, these cases suit nicely the purpose of illustrating our claim about the interplay between individual heterogeneity and aggregate persistence. Besides, it gives us an idea of how the dollar deposit market is shared among various types of depositors.

In section 3 we develop a stylized model where agents face noisy information and differ in their ability to forecast when taking portfolio decisions. An important result from this setup is that the dynamics of the individual's optimal portfolio depends on her prediction errors of future dollar returns. It turns out that it is optimal for agents to be cautious when modifying the currency composition of their deposits as there is uncertainty on the quality of the data agents receive. This caution is reflected in portfolios that may adjust in a relatively slow fashion. Finally, we show that upon aggregation of the individual dollarization decisions it is possible to generate a very persistent economy-wide dollarization ratio.[11]

In section 4 we test the empirical hypotheses of the theoretical model and find supportive evidence from aggregate data of three Latin American countries and Poland. Particularly, the results suggest that the distributions of "forecasting abilities" behind the aggregate dollarization ratios are very spread. We regard this result as consistent with the idea of financial experts sharing the dollar market with blacksmiths that save in dollars. In section 5 we reformulate the theoretical model in order to reflect a different source of informational heterogeneity, and provide further empirical evidence.

Section 6 concludes and gives some policy recommendations. Derivations and complementary results are shown in the appendix.

[10] The figures used in section 2 come from the Central Bank of Peru and the National Bank of Poland. The facts discussed there are recorded in the annual reports of these institutions.

[11] Our approach is related to other branches of the literature. For instance, Lewbel (1994) uses aggregate information to test heterogeneity on consumption dynamics whereas Michelacci (2004) explains the high degree of persistence of output with the cross-sectional heterogeneity of productive firms.

2 Two illustrative cases

As documented by Savastano (1996), dollarization emerges progressively in response to macroeconomic instability, particularly high levels of inflation, showing a well-defined pattern: first agents replace domestic currency as reserve of value, holding usually dollars outside the financial system ("under the mattress"). Then, the dollar is used in some transactions, typically involving real estates and durable goods, and eventually some prices are set in dollars. Most governments later on allow banks to issue deposits in foreign currency to avoid financial disintermediation.[12] The actual experience of various countries shows that within a year an economy can increase its dollarization ratio enormously, see Figures 1(a) and 2(a).

On the other side, episodes of *de*dollarization (i.e., a sustained reduction in the dollarization ratio) are not very common and thus there is no well-established pattern in the literature. Yet, if ever happened, the dedollarization process is likely to be slow. The analysis of these events, as opposed to the increase of dollarization, provide very useful information about the way different depositors decide the currency composition of their savings and on how they respond to news coming from the macroeconomic environment.

2.1 Peru in the early 2000's

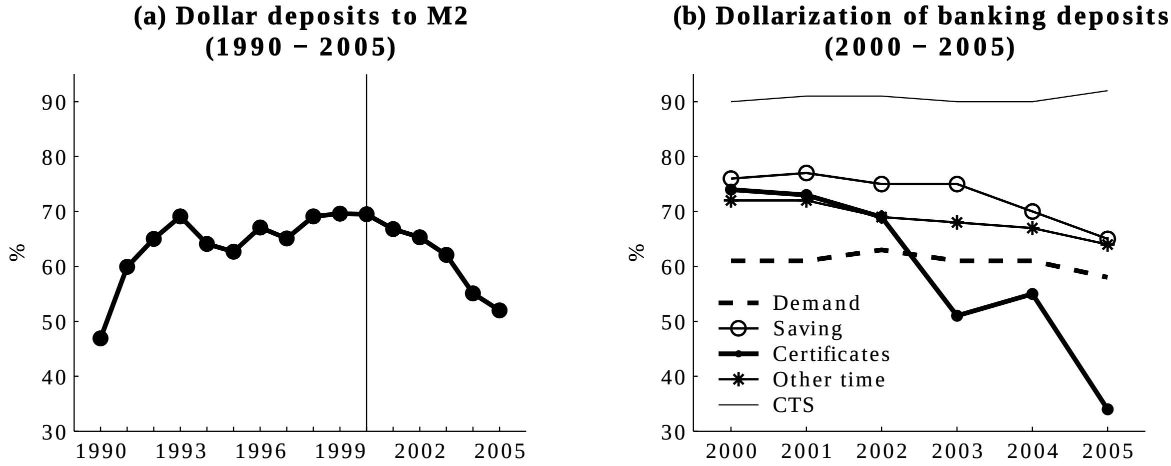
Although the Peruvian dollarization experience shares various of the aforementioned features, it has its own appeal.[13] As shown in Figure 1(a), in 1991 (after a four-digit hyperinflation in 1990) the ratio was 60% and has remained fluctuating roughly between 65% and 70% for a decade. Since 2000, it has shown a sustained reduction to about 50% in 2005. Of course, 50% is still a big number, but there are some interesting facts behind this recent drop.

There are at least two forces driving this decrease. Firstly, after 8 years of announcing inflation targets within a monetary targeting regime (since 1994) and after 5 years of having achieved a one-digit inflation rate, the Central Bank announced the adoption of a fully fledged inflation targeting regime in 2002. This has helped to anchor inflation expectations and has reduced inflation and nominal interest rate volatility. Secondly, between 2001 and 2005, the nominal and real exchange rates have appreciated (6.2% and 5.1%) as a result of a very favorable foreign environment: increasing terms of trade leading to an export boom and very low international interest rates. In a nutshell, the real return to holding deposits dollars vis-à-vis holding deposits in domestic currency has fallen considerably in the early 2000's.

Figure 1(b) shows deposit dollarization by type of deposit: demand, savings and a breakdown of time deposits in certificates, "CTS" and others. A glimpse of the figure reveals that both demand and "CTS" deposits have not reacted to the recent change in the dollar real return

[12] See also Kamin and Ericsson (2003), De Nicoló *et. al* (2005) and Levy Yeyati (2006).
[13] See Quispe (2000) for a careful historical account of the dollarization experience in Peru.

Figure 1. Deposit Dollarization in Peru



Source: Central Bank of Peru.

trend. Demand deposits accounts for about 20% of total deposits and as the most liquid, almost transactional kind of deposit the flat pattern is justified. On the other side, the CTS is the Peruvian version of an unemployment insurance; by law, it is hold exclusively by individuals and can be claimed only when an individual becomes unemployed. The CTS deposits have reacted even less than the demand deposits, which is puzzling.

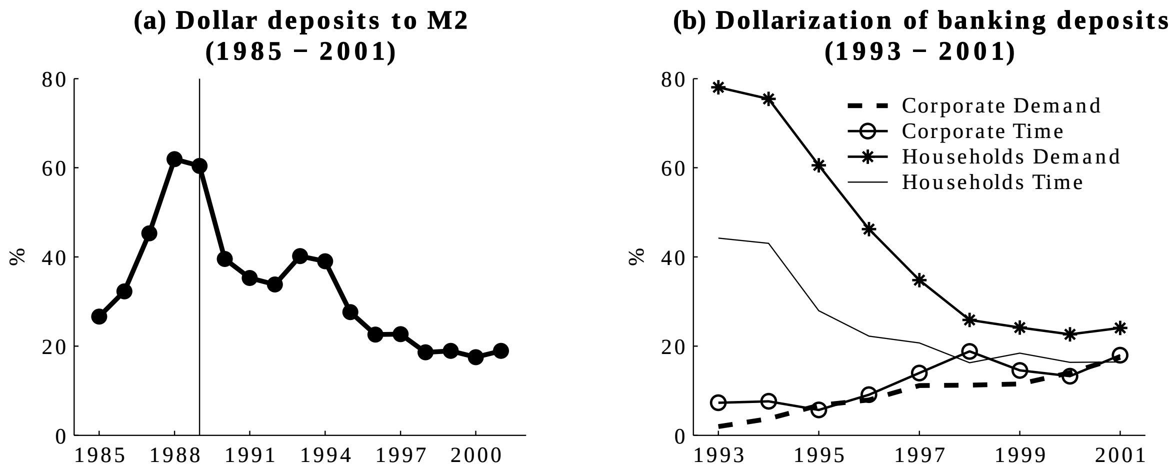
The figure also shows a moderate downward trend in the savings and other time deposits. About 80% of the saving and roughly half of the other time deposits are held by individuals. From 2001 to 2005 both ratios have decreased in about 10%. What is remarkable from Figure 1(b) is the strong reaction of the certificate of deposits ratio which has fallen in almost 40%, and with no doubts is driving the fall in the aggregate ratio of Figure 1(a). The interesting fact is that although the certificate of deposits have similar term than the CTS and the other time deposits, they are mainly held by firms and not individuals.

2.2 Poland towards a market economy

The Polish experience is regarded as the most successful shift from a planned to a market-oriented economy, and is a thriving example of dedollarization. By the end 1980's, Poland was on the verge of a profound economic crisis. The huge distortions on relative prices and the cumulative fiscal deficits, inherited from the years of central planning, induced a rapid increase in inflation that reached its historical maximum of 550% in 1989. In response to this unstable macroeconomic environment, dollarization ratios increased rapidly, from levels around 20% in 1985 to a peak of 60% in 1989. This is shown in Figure 2(a).

After the introduction of a series of pro-market reforms and of a stabilization program

Figure 2. Deposit Dollarization in Poland



Source: National Bank of Poland.

(the so-called "shock-therapy"),[14] dollarization ratios dropped to averages of 40% percent by the end of 1993, hand-to-hand with the reduction of inflation (from 500% to 36%). As the macroeconomic conditions kept improving, additional institutional reforms were put in place. Notably, in 1997 the National Bank of Poland was granted independence and a well-defined objective: to guarantee price stability. Dollarization decreased even more reaching by 2001 the level of 18%, comparable with that of developed European economies, as the UK.

A common feature of the Polish experience with the Peruvian one discussed above is the observed heterogeneity of dollarization dynamics among type of deposits. Figure 2(b) reveals that by the end of 1993, the difference between the dollarization ratios of households and firms was of the order of 70% for time deposit and 40% for demand deposits. These differences remained on the range of 20% for more than 4 years.

2.3 Moral

The differences between how individuals and firms decide their portfolio composition are obvious. Usually firms have more resources allocated to the management of their funds, whereas individuals often base their decisions on their experience, those of some neighbors and their limited access to information. Moreover, the decision-making even within firms or within individuals is likely to be dissimilar. Our brief inspection of the Peruvian and Polish experiences illustrates our main claim that these differences accounts for much heterogeneity in dollarization decisions. We next analyze how this translates into persistence.

[14] A drastic series of institutional and market reforms were put in place in 1990: the government liberalized controls of almost all prices, eliminated most subsidies, abolished administrative allocation of resources in favor of trade, promoted free establishment of private businesses, liberalized the system of international economic relations, and introduced an internal currency convertibility with a currency devaluation of 32%.

3 A simple model

We use a simple framework to show how the combination of imperfect, noisy information on real returns of foreign assets, and specially the heterogeneity among market participants can generate a persistent degree of dollarization.

The model economy is populated by a number of almost identical individuals. They have the same endowment, which is normalized to one, and the same preferences, but they differ in their ability to process information and therefore in their expectations on future outcomes.[15]

Every period agents choose the composition of their portfolio between two assets, one that offers a fixed real return R^P which is denominated in domestic currency (*peso* from now on) and the other denominated in dollars with real return R_t^D. For the sake of concreteness we will focus on the real excess of return of the dollar over the peso asset,

$$R_t = R_t^D - R^P \tag{1}$$

3.1 Portfolio decision

Depositors are risk adverse. Individual i devotes an amount x_{it} of her savings to the dollar asset and the remaining $1 - x_{it}$ to purchase the asset in pesos. We follow Ize and Levy Yeyati (2003) in assuming a simple mean-variance utility function. Since the portfolio decision is ex-ante and based on imperfect information on real returns, the utility for individual i is defined in terms of the conditional expectation for period $t+1$ with information up to period t,[16]

$$\begin{aligned} U_{it} &= E_t\left[x_{it}R_{t+1}^D + (1-x_{it})R^P\right] - \left(\tfrac{1}{2}\right)\text{var}_t\left[x_{it}R_{t+1}^D + (1-x_{it})R^P\right] \\ &= E_t\left[x_{it}R_{t+1} + R^P\right] - \left(\tfrac{1}{2}\right)\text{var}_t\left[x_{it}R_{t+1}\right] = x_{it}\hat{r}_{it+1} + R^P - \left(\tfrac{1}{2}\right)(x_{it})^2 \text{v}_{it+1} \end{aligned} \tag{2}$$

where $\hat{r}_{it+1}$ and v_{it+1} are the mean and variance of the excess return R_t that individual i expects for period $t+1$, conditional on information up to period t.

The value of x_{it} that maximizes (2) is

$$x_{it} = \frac{\hat{r}_{it+1}}{\text{v}_{it+1}} \tag{3}$$

Thus, agents will increase their dollar deposits when they expect a higher real return on this asset for the same expected variance, or when the expect a lower variance for a given level of excess of return.

[15] Our analysis holds for agents with heterogenous endowments, i.e. wealth/income inequality, as long as they are correlated with the abilities to process information. See appendix A for details.

[16] We have imposed a value of one to the risk aversion parameter in the utility function. This assumption is harmless to our results.

3.2 Forecasting

In period t, the excess return R_t cannot be perfectly observed. What is publicly known is a noise-ridden version of R_t, $S_t = R_t + \varepsilon_t$ where $\varepsilon_t \sim iid(0, \sigma^2)$. In practice, these variables can be understood as ex-ante and ex-post differentials, respectively. As equation (3) reveals, the only relevant pieces of information for the portfolio decision are the excess return and its variance. To make things easier, we postulate that each agent focuses directly on forecasting R_t, and not on forecasting its components (R_t^D or R^P, which may imply forecasting inflation, depreciation, confiscation risk and so on).

Each individual has a *forecasting model* of the form

$$\begin{array}{lllr} r_{it+1} & = & \mu(1-a_i) + a_i r_{it} + w_{it+1} & w_{it} \sim iid(0, \sigma_w^2) \\ S_t & = & r_{it} + \epsilon_{it} & \epsilon_{it} \sim iid(0, \sigma_\epsilon^2) \end{array} \tag{4}$$

Since S_t is a noisy indicator, individual i has first to extract the *signal* r_{it} (i.e., "nowcasting") and then forecast its mean and variance to implement (3).

As it can be seen from (4), each individual is given a value of $|a_i| \leq 1$ drawn from the distribution $F(a)$, to perform her predictions, and this value alone determines the whole forecasting model. This is the only source of (cross-sectional) heterogeneity in this setup. The parameter μ is common to all individuals, indicating that they share the same long-run forecast. The variances of the disturbances, σ_w^2 and σ_ϵ^2, are also assumed to be the same across individuals.

To be precise, what differentiates the individuals is the speed with which they adjust their short-run forecasts as new information becomes available. We interpret this heterogeneity as differences in the ability people have to forecast,[17] which translates directly to portfolio differences among individuals.

Define $\mathrm{v}_{it} = E[(r_{it} - \hat{r}_{it})^2]$ as the mean squared error (MSE) of the predictor $\hat{r}_{it}$. Standard results from the signal extraction literature lead us to the optimal prediction rule[18]

$$\hat{r}_{it+1} = \mu(1-a) + a\,\hat{r}_{it} + \mathrm{k}_{it}(S_t - \hat{r}_{it}) = \mu(1-a) + a\,\hat{r}_{it} + \mathrm{k}_{it}\xi_{it} \tag{5}$$

where the forecasted value of r_{it} for next period is the projection of today's forecasted value plus a correction, an updating $\xi_{it} = \epsilon_{it} + (r_{it} - \hat{r}_{it})$, that is proportional to the latest prediction error incurred. The value of k_{it} is given by the (adjusted) ratio of the MSE of $\hat{r}_{it}$ to the variance

[17] For instance, if the true DGP of R_t is a random walk, then the predictions of those with $a_i \simeq 1$ will outperformed that of the others (the true DGP of R_t is irrelevant in deriving our main results).

[18] The reader that is familiar with state-space modelling will note that the recursions (5) and (7) below are straightforward applications of the Kalman filter. See Ljungqvist and Sargent (2000, ch. 2 and ch. 21) and Harvey and De Rossi (2006) for further details.

of the noisy indicator,

$$\mathrm{k}_{it} = a_i \left(\frac{\mathrm{v}_{it}}{\mathrm{v}_{it} + \sigma_\epsilon^2} \right) \tag{6}$$

The MSE of $\hat{r}_{it}$ evolves according to the following recursion

$$\mathrm{v}_{it+1} = \frac{\mathrm{v}_{it}(a_i^2 \sigma_\epsilon^2 + \sigma_w^2) + \sigma_\epsilon^2 \sigma_w^2}{\mathrm{v}_{it} + \sigma_\epsilon^2} \tag{7}$$

It is clear from equation (7) that $\mathrm{v}_{i\tau+1} = f(\mathrm{v}_{i\tau})$. There is a fixed point such that $\mathrm{v}_i = f(\mathrm{v}_i)$[19] and moreover, since $f'(\mathrm{v}_i) < 1$ it is globally stable: regardless of the initial condition v_{i0} we have that $\mathrm{v}_{i\tau} \to \mathrm{v}_i$ and consequently $\mathrm{k}_{i\tau} \to \mathrm{k}_i = a_i \mathrm{v}_i (\mathrm{v}_i + \sigma_\epsilon^2)^{-1}$ as $\tau \to \infty$. This means that as τ becomes larger, i.e. as each individual has performed the signal extraction exercise a number of times, the updating process defined in (5) and (7) converges to an equilibrium rule.[20] If we assume that this learning process was initialized long before period t then we can safely treat v_{it} and k_{it} as constants. This fact simplifies the calculations considerably without compromising our conclusions.

To have a better grasp of the way heterogeneity among agents affects their forecasts (and portfolios), assume for a moment that $\mu = 0$ and solve (5) recursively to get

$$\hat{r}_{it+1} = \mathrm{k}_i \sum_{j=0}^{\infty} (a_i - \mathrm{k}_i)^j S_{t-j}$$

It is clear from this geometrically distributed lag expression that different draws of a are associated with different ways of weighting the available information (the noisy indicators up to period t) in order to produce a forecast.[21]

3.3 Individual dynamics

Using the fact that $\mathrm{v}_{it} \to \mathrm{v}_i$, $\mathrm{k}_{it} \to \mathrm{k}_i$ and the optimal updating/forecasting rule (5), the optimal dollar investment (3) boils down to

$$x_{it} = \frac{\hat{r}_{it+1}}{\mathrm{v}_{it+1}} = \frac{\hat{r}_{it+1}}{\mathrm{v}_i} = \frac{\mu(1 - a_i) + a_i \hat{r}_{it} + \mathrm{k}_i \xi_{it}}{\mathrm{v}_i} = a_i x_{it-1} + \frac{\mu(1 - a_i)}{\mathrm{v}_i} + \left(\frac{\mathrm{k}_i}{\mathrm{v}_i} \right) \xi_{it} \tag{8}$$

[19] The fixed point is the positive root of $\mathrm{v}_i^2 + [(1 - a_i^2)\sigma_\epsilon^2 - \sigma_w^2]\mathrm{v}_i - \sigma_\epsilon^2 \sigma_w^2 = 0$.

[20] Convergence is monotonic ($\mathrm{v}_{i\tau} \geq \mathrm{v}_{i\tau+1} \geq \mathrm{v}_i$) because $\mathrm{v}_{i\tau+1}$ is based on more information than $\mathrm{v}_{i\tau}$. This updating scheme is *E-stable*, in the sense of Evans and Honkapohja (2001) in adaptive learning.

[21] As noted in Harvey (1989, ch. 4), the forecasting model converges to the popular Exponential Smoothing method (ES) as $a \to 1$. However, the scheme explained here is optimal in the sense that it minimizes the one step ahead MSE, whereas ES is basically *ad hoc*. We explore this issue further in section 5.

After plugging (6) into (8), we get

$$x_{it} = a_i x_{it-1} + \frac{\mu(1-a_i)}{\mathrm{v}_i} + a_i \left(\frac{1}{\mathrm{v}_i + \sigma_\epsilon^2} \right) \xi_{it} \tag{9}$$

The individual's dollarization ratio follows a simple AR(1) process. As such, it exhibits some degree of persistence which depends on the value of a_i. Note that a shock in ξ_{it} – an informational update, a news – changes x_{it} on impact, as it changes individual i's forecast of R_t and makes her revising her optimal portfolio accordingly. If $\xi_{it} > 0$ ($\xi_{it} < 0$), x_{it} increases (decreases) as the expected excess of return for holding dollars is higher (lower). *Ceteris paribus*, in subsequent periods the expected path of R_t smoothly adjusts towards its long-term value, and so does the dollar share in the portfolio.

The dynamics of individual dollarization decisions reflects the fact that with noisy signals of returns, individuals have to rely on past information to optimally forecast them, and have to react with caution to news. To the extent that past portfolio decisions contain past information of returns, it becomes optimal for individuals to make their dollarization ratios depended on past dollarization ratios.[22]

3.4 Aggregate dynamics

In a static world the effects of aggregation are well-known: it tends to smooth away individual erratic movements and to fill in discontinuities that may be present at the disaggregate level. Within a dynamic framework, aggregation also increases persistence.[23] To see why consider a group of individuals who hold a small amount of the dollar asset and face an aggregate shock that makes it more attractive (e.g., a strong real depreciation). According to (9), these individuals will increase their dollar holdings immediately. But then, this group will also revise their expectations about future returns in favor of the dollar asset, thereby perpetuating the impact effect of the shock on aggregate dollarization. Thus, the moderate persistence in the individual portfolio formation due to the lack of perfect information, summarized in equation (9), is exacerbated by aggregation.[24]

Let X_t be the economy-wide dollarization ratio. In Appendix B it is shown that aggregation of (9) across the distribution of a renders the following process

$$X_t = \sum_{r=1}^{\infty} A_r X_{t-r} + \tilde{M} + \tilde{U}_t \tag{10}$$

[22] A similar result but in a different setup can be found in Aoki (2003). In that paper the central bank sets interest rates in an environment with noisy information on output and inflation. The optimal policy rule implies some persistence coming from the cautiousness that the lack of perfect information demands.

[23] The classic reference for the econometrics of this effect is Granger (1980), which assumes that $F(a)$ is a Beta distribution. See also Pesaran (2003) and Zaffaroni (2004) for recent developments.

[24] See Michelacci (2004) for a similar analysis.

where the A_s ($s = 1, 2, \ldots$) are coefficients, $\tilde{M}$ is a constant and $\tilde{U}_t$ is an aggregate serially uncorrelated disturbance. As suggested before, the remarkable fact is that although at the individual level the dollar share in the portfolio follows an AR(1) process, it becomes AR(∞) at the aggregate – usually known as a process exhibiting *long-memory*.

As stressed by Lewbel (1994), the coefficients in (10) are tightly related to the shape of $F(a)$. In Appendix B it is also shown that they satisfy the recursion

$$A_s = m_s - \sum_{r=1}^{s-1} m_{s-r} A_r \qquad (s = 1, 2, \ldots) \tag{11}$$

where m_s is the s-th *moment* of the distribution of a, $m_s = \int a^s \, \mathrm{d}F(a)$. Hence, it is easy to verify that

$$\begin{aligned} \text{mean}(a) &= m_1 = A_1 \\ \text{variance}(a) &= m_2 - m_1^2 = A_2 \\ \text{skewness}(a) &= (m_3 - 3m_1 m_2 + 2m_1^3)(m_2 - m_1^2)^{-3/2} = (A_3 - A_1 A_2)(A_2)^{-3/2} \end{aligned}$$

These relations allow us to determine how the distribution of forecasting abilities affects persistence at the aggregate level. The higher A_1, the higher the mean which implies that the *average* individual has herself a more persistent behavior, rendering subsequently a more persistent X_t. On the other side and strikingly, a higher A_2 renders also more persistence: the higher the heterogeneity among individuals, the more persistent the aggregate dollarization ratio. Finally, as pointed out by Zaffaroni (2004), the low frequency behavior of the aggregate is determined by the shape of the cross sectional distribution as $a \to 1^-$. Hence, a distribution with a heavy left tail ($A_3 < A_1 A_2$), which indicates a higher mass of persistent individuals ($a \approx 1$), would suggest higher aggregate persistence.

It is now clear that this framework can be tested straightforwardly. If the estimates of A_s using aggregate data are inconsistent with the notion of various dynamic processes that have been aggregated into (10), then we are to reject the model.[25] The most obvious symptoms of contradiction would be a non-positive estimate of A_2, the variance of $F(a)$,[26] or a very negative value for A_1, the mean.

It is important to bear in mind that the amount of information about individual behavior that can be inferred from aggregate data is unquestionably limited. Different assumptions regarding individual decisions can be found to be consistent with a given observed aggregate variable. Yet, there are some facts reported below that are supportive to the main hypothesis of this paper and the predictions of the theoretical model.

[25] Or the assumptions behind the aggregation, see Appendix B.

[26] Note that $A_2 = 0$ implies a degenerate distribution of a on the point A_1, i.e. a model with a representative agent or identical individuals.

4 Empirical evidence

In this section we test whether the dynamics of the aggregate dollarization ratio in selected countries can be regarded as coming from the aggregation of heterogeneous depositors. In other words, we estimate the parameters A_s in equation (10) and investigate, from the estimated moments of the underlying distribution $F(a)$, the extent of heterogeneity among participants in the dollar deposit market.

4.1 Econometric issues and data

Three points are worth mentioning before presenting some results. Firstly and unsurprisingly every dollarization ratio X_t we considered has a unit root[27] and to avoid well-known biases in the estimation of autoregressive coefficients when a unit root is present we estimate (10) in first differences,

$$\Delta X_t = \sum_{r=1}^{\infty} A_r \Delta X_{t-r} + U_t^{\dagger} \tag{12}$$

Appendix B shows that (12) is not only the first-differenced version of (10), but is also the result of aggregating (9) after first-differentiating. Hence, the coefficients in (12) are indeed the same as in (10). The disturbance $U_t^{\dagger}$ is autocorrelated and heteroscedastic[28] so robust inference is required.

Secondly, due to data limitations it is not possible to estimate equation (12) as it stands. Data are finite, so a truncation in the lags of the $AR(\infty)$ process is unavoidable.

Lastly, if convenient, we consider even richer dynamics than the suggested by our very stylized theoretical model by introducing a MA(1) component in (12). In practice, this fact has no other implication for our analysis than to produce better estimates of the A_s. As noted by Lewbel (1994), with a MA component present only a finite number of the moments of $F(a)$ can be recovered as an infinite autoregression in X_t (or in ΔX_t) cannot be separated from the MA parameter, say θ. This is a theoretical rather than empirically substantive concern; as noted earlier, our attempt is not to recover every moment of $F(a)$, but just the first few.

We gathered information for Peru and Uruguay (two highly dollarized countries), Mexico and Poland. Data are quarterly spanning roughly from the mid-1980's to the mid-2000's. As it is customary in the dollarization literature, X_t is measured as the ratio of foreign currency deposits from the private sector in the domestic banking system to M2.[29] This information

[27] Results of unit root tests are available upon request to the authors. See also Appendix C.

[28] See Pesaran (2003) for further details.

[29] A popular alternative definition of the dollarization ratio discriminate between residents and non-residents, which includes deposits by residents abroad (Ize and Levy Yeyati, 2003). We did not include this definition in our empirical work as the corresponding available time series are shorter for the pool of countries analyzed.

Table 1. ARIMA models of the deposit dollarization ratio in selected countries

ARIMA model	A_1	A_2	A_3	A_4	θ	$A_3 - A_1A_2$	$\bar{R}^2$
Mexico (1985.Q4 to 2005.Q3, $N = 77$)							
$(4, 1, 0)$	0.221*	0.199*	−0.192*	0.114**		−0.236*	0.221
	(0.078)	(0.078)	(0.072)	(0.064)		(0.095)	
$(4, 1, 1)^\star$	0.480*	0.195*	−0.216*	0.251*	−0.097*	−0.310*	0.261
	(0.111)	(0.094)	(0.063)	(0.047)	(0.018)	(0.086)	
Peru (1980.Q1 to 2005.Q3, $N = 94$)							
$(2, 1, 0)^\star$	0.173*	0.142*				−0.024**	0.200
	(0.063)	(0.058)				(0.013)	
$(2, 1, 1)$	0.186**	0.139*			−0.058	−0.026	0.173
	(0.094)	(0.065)			(0.143)	(0.016)	
Poland (1985.Q4 to 2002.Q4, $N = 69$)							
$(2, 1, 0)^\star$	0.474*	0.113*				−0.053*	0.215
	(0.016)	(0.052)				(0.024)	
$(2, 1, 1)$	0.476*	0.111*			−0.007	−0.053*	0.275
	(0.010)	(0.049)			(0.045)	(0.024)	
Uruguay (1985.Q1 to 2005.Q3, $N = 83$)							
$(2, 1, 0)$	0.218*	0.290*				−0.063*	0.153
	(0.091)	(0.116)				(0.029)	
$(2, 1, 1)^\star$	0.265**	0.215*			−0.093*	−0.057**	0.196
	(0.147)	(0.055)			(0.034)	(0.033)	

Maximum likelihood estimates. Figures in parentheses are robust (consistent) standard errors. * [**] denotes significance at a 5% [10%] level. The standard error of the third central moment $A_3 - A_1A_2$ was computed with the delta method. $\bar{R}^2$ is the adjusted R^2. Regressions include a constant and, if necessary, a few dummy variables for outlier removal. In all reported equations, Breusch-Godfrey and Jarque-Bera tests suggested uncorrelated and normally distributed residuals. The preferred specifications are marked with a $\star$.

is widely available and our sources are the websites of the various central banks and the International Financial Statistics database, IFS. The regression with the shortest time series (Poland) has $N = 69$ observations; the one with the largest (Peru), $N = 94$.

4.2 Results

For each country an ARIMA(2,1,0) was first fitted. Then, we test for residual autocorrelation and include further lags until the residuals appear serially uncorrelated. In every case, no more than 2 lags is needed, but in Mexico when the lag length is 4. For robustness sake we then include a MA component in the best autoregressive specification. Table 1 reports for each country the best autoregressive model, ARIMA(2,1,1) or ARIMA(4,1,0), and the corresponding ARIMA(2,1,1) or ARIMA(4,1,1) equations. The column labelled θ contains the estimated MA coefficient. For each country we have marked our preferred specification, i.e. the more parsimonious model that describes the data sufficiently well, with a $\star$.

A remarkable fact from Table 1 is that the estimates for Peru are close to those of Uruguay, whereas the Mexican estimates are similar to the Polish. Recall that Peru and Uruguay are heavily dollarized (above 50%), whereas Mexico and Poland, even though have reported sizeable dollarization ratios by the early or mid-90's, have dollarization ratios less than 30%. There are however, common features among all 4 countries and we will focused on those next.

A finding that is robust among countries and specifications within the same country, is that the coefficients A_1 and A_2 are significantly positive. Recall that A_1 is the mean of $F(a)$, which does not appear to be particularly close to one in any case, and A_2 is its variance. Moreover, in Peru and Uruguay the coefficients are of comparable magnitude, $A_2 \approx A_1$, which means that the underlying $F(a)$ is very spread, the a's are fairly heterogeneous. From the preferred model for Peru we have that $A_1 \pm 2\sqrt{A_2} \approx [-0.60, 0.92]$, whereas in the Uruguayan case this interval is even wider, $A_1 \pm 2\sqrt{A_2} \approx [-0.66, 1.20]$. For the Mexican case, $A_1 \pm 2\sqrt{A_2} \approx [-0.40, 1.36]$ whereas for the Polish, $A_1 \pm 2\sqrt{A_2} \approx [-0.20, 1.15]$. These estimates imply coefficient of variations $\sqrt{A_2}/A_1$ of 2.18 for Peru, 1.75 for Uruguay, 0.91 for Mexico and 0.71 for Poland. Hence, the highly dollarized economies appear to have a spreader $F(a)$ which is consistent with the idea of decreasing participation costs as dollarization expands. Yet, even in Mexico and Poland (where $A_1 > A_2$) the underlying heterogeneity is estimated to be high.

From the above intervals it can be seen that the largest mass of depositors has a stationary behavior, $|a| < 1$. A more controversial finding is that some individuals seem to have explosive dynamics, $a > 1$. This is, nonetheless, not to be interpreted narrowly as the *same* depositors having $a > 1$ all time periods, but as a mass of individuals having this sort of behavior from time to time. The same argument applies to the fact that some depositors appear to have $a < 0$, which implies an erratic, saw-shaped, individual dollarization pattern.

The estimates of the implied third central moment $A_3 - A_1 A_2$ in each country suggest that $F(a)$ is skewed to the left. Provided that $A_1 > 0$, a left-skewed $F(a)$ would be expected if it were the mixture of a mass point above the mean (relatively persistent individuals, those who change their portfolio slowly) and some individuals with a close to zero (corresponding to those who change their portfolio quickly). Negative skewness, thus, is consistent with a financial expert sharing the dollar market with a non-expert blacksmith saving in dollars.

5 A reinterpretation

To check the robustness of our results, in this section we slightly modify our theoretical model. In section 3, the differences among agents were centered on how fast each adjusts her forecast (and hence her portfolio) as news become available. We dubbed this differences as coming from different abilities to forecast. Next, we focused on heterogeneity in the abilities people have to extract useful signals. We postulate that individuals face idiosyncratic signal-to-noise ratios, q_i. This rationalizes in a simple manner the fact that those with high q_i, the financial experts, are able to extract more information from the noisy indicator S_t than those with low q_i, the blacksmiths. In contrast to the blacksmiths, the financial experts might be able to distinguish whether changes in S_t reveal changes in R_t or are just due to noise.

5.1 Reformulating the model

We collapse system (4) to what is known as a *local level model*,[30]

$$\begin{aligned} r_{it+1} &= r_{it} + w_{it+1} \qquad w_{it} \sim iid(0, \sigma_w^2) \\ S_t &= r_{it} + \epsilon_{it} \qquad \epsilon_{it} \sim iid(0, \sigma_{\epsilon i}^2) \end{aligned} \tag{13}$$

and now we allow the variance of the noise to vary, $\sigma_{\epsilon i}^2$. The signal-to-noise ratio is $q_i = \sigma_w^2/\sigma_{\epsilon i}^2$ and plays a key role in determining how the noisy observations are weighted for signal extraction (and prediction). The higher is q_i the more past observations are discounted in forecasting the future.

For expositional convenience we define $\tilde{\mathrm{v}}_{it} = \mathrm{v}_{it}\sigma_{\epsilon i}^{-2}$. The updating scheme of the MSE of $\hat{r}_{it}$ – formerly equation (7) – and its equilibrium value become

$$\tilde{\mathrm{v}}_{it+1} = \frac{\tilde{\mathrm{v}}_{it}(1+q_i)+q_i}{\tilde{\mathrm{v}}_{it}+1} \qquad \text{and} \qquad \tilde{\mathrm{v}}_i = \frac{q_i + \sqrt{q_i^2+4q_i}}{2} \tag{14}$$

whereas the forecasting rule – (5) before – can be written now as

$$\hat{r}_{it+1} = (1-\mathrm{k}_i)\,\hat{r}_{it} + \mathrm{k}_i S_t \qquad \text{where} \qquad \mathrm{k}_i = \frac{\tilde{\mathrm{v}}_i}{\tilde{\mathrm{v}}_i+1} \tag{15}$$

Given (15) the optimal dollar investment follows the proccess

$$x_{it} = (1-\mathrm{k}_i)\,x_{it-1} + \left(\frac{1}{\tilde{\mathrm{v}}_i+1}\right) S_t \tag{16}$$

It is easy to show that k_i is increasing in q_i,[31] which implies that those individuals with high q_i gain more information from the signal each period. More fundamentally, this also implies that those individuals will have less persistent dollarization ratios. As equation (16) shows, the higher the k_i, the lower the degree of persistence of dollarization ratios. Note as well that a change on the signal S_t changes the dollarization ratio on impact, with higher responsiveness from individuals endowed with a higher ability to extract information.

[30] See Harvey (1989, ch. 4) and footnote 21.

[31] Since $\partial\mathrm{k}_i/\partial q_i = \partial\mathrm{k}_i/\partial\tilde{\mathrm{v}}_i \cdot \partial\tilde{\mathrm{v}}_i/\partial q_i$ with $\partial\mathrm{k}_i/\partial\tilde{\mathrm{v}}_i = (\tilde{\mathrm{v}}_i+1)^{-2} > 0$ and $\partial\tilde{\mathrm{v}}_i/\partial q_i = 0.5+(q_i+2)(q_i^2+4q_i)^{-0.5} > 0$.

5.2 Aggregation and further econometric issues

With a slight abuse of notation, call $a_i = 1 - \mathrm{k}_i$. Recall now that $S_t = R_t + \varepsilon_t$, where ε_t is an aggregate shock. Then, the aggregation of (16) (see Appendix B) leads to

$$X_t = \sum_{r=1}^{\infty} A_r X_{t-r} + \sum_{r=0}^{\infty} \beta_r R_{t-r} + \hat{U}_t \tag{17}$$

which as opposed to (10) includes a distributed lag of R_t. This difference is the consequence of postulating different sources of heterogeneity at the individual level, and has a direct implication for our analysis so far: if the individual heterogeneity is best approximated by (13) rather than (4), then the estimates of Table 1 may be biased due to the omission of relevant variables. Next, we augment the ARIMA models of Table 1 to investigate whether this presumed omission changes our main conclusions.

As discussed in section 4.1, the actual object to be estimated is

$$\Delta X_t = \sum_{r=1}^{p_X} A_r \Delta X_{t-r} + \sum_{r=0}^{p_R} \beta_r \Delta R_{t-r} + U_t^{\ddagger} \tag{18}$$

where p_X and p_R are finite lag lengths. The presence of R_t and its lags in (18) follows directly from the fact that the individuals in the theoretical model base their decisions exclusively on this variable. Nonetheless, a richer modelling framework can easily extend (18) to

$$\Delta X_t = \sum_{r=1}^{p_X} A_r \Delta X_{t-r} + \sum_{r=0}^{p_R} \beta_r^D \Delta R_{t-r}^D + \sum_{r=0}^{p_R} \beta_r^P \Delta R_{t-r}^P + U_t^{\ddagger} \tag{19}$$

As $R_t = R_t^D - R_t^P$, equation (19) encompasses (18) which is a restricted version with $\beta_r^D = -\beta_r^P$ for every r. For this reason, we will focus on (19) from now on.

An empirical issue that raises with the introduction of the real returns in the aggregate equations is, precisely, how to measure them. The "true" returns involve expectations of future macroeconomic variables, which historical data are barely available for the countries in our analysis. Call i_t^P and i_t^D the nominal interest rates in domestic currency and US dollars, respectively, δ_t the nominal depreciation (i.e., the percent change of the nominal exchange rate, domestic currency per US dollar) and π_t the CPI inflation. We entertain two measurements of the real returns:

ex-ante: $$R_t^P = \frac{1 + i_t^P}{1 + \pi_{t+1}} - 1 \qquad R_t^D = \frac{(1 + i_t^D)(1 + \delta_{t+1})}{1 + \pi_{t+1}} - 1$$

ex-post: $$R_t^P = \frac{1 + i_t^P}{1 + \pi_t} - 1 \qquad R_t^D = \frac{(1 + i_t^D)(1 + \delta_t)}{1 + \pi_t} - 1$$

CPI and nominal exchange data are readily available. For i_t^P we use the deposit rate in domestic currency for Peru, Poland and Uruguay and the saving rate in domestic currency for Mexico. For i_t^D, we found data on the interest rate paid to domestic deposits in dollars only in the case of Peru and Uruguay. For Mexico and Poland we approximate i_t^D with the deposit rate in the US.[32] Our sources are still the central banks and the IFS.

Finally, the presence of a contemporaneous return (19) may rise the possibility of endogeneity bias. We use a 2SLS procedure to estimate this equation. The instruments are listed in the note to Table 2. It is worth mentioning that OLS or the exclusion of the contemporaneous returns did not alter the main results of this robustness check.[33]

5.3 Results

Table 2 displays the estimation results. To save space we do not report the coefficients of the returns (as they are not of direct interest for our analysis) but do report an F-statistic assessing its overall significance. We set the lag length $p_X = 3$. This is the best choice for Mexico; for the other countries, the optimal is $p_X = 2$, but we still set $p_X = 3$ to ensure that no autoregressive effect is ignored. The choice of p_R, reported in the table, responds to the minimization of the Schwarz criterion.

Recall that by estimating the augmented equations we are assessing whether the results of Table 1 are robust. So, are they robust? In general they are. A quick comparison of the estimates in Table 2 with those in Table 1 reveals that due to the presence of the returns, the fit of the various equations increases, but the estimates of A_1, A_2 and $A_3 - A_1A_2$ do not change much. The notable exception to this pattern is the Mexican case when the returns are measured in the *ex-post* manner, as A_1 losses statistical significance. However, the main claim of the previous sections still holds, qualitatively and almost quantitative: the heterogeneity of decision-makers that underlies the aggregate dollarization ratios is high, and this fact leads to aggregate dollarization persistence.

6 Concluding remarks

In countries with high dollarization ratios, participation in the dollar deposit market has become massive. Financial deregulation, liberalization, innovation and informal currency

[32] Unfortunately we could not find time series long enough of country risk to have a better measure of R_t^D in these two countries. The estimation results, though, were robust when we considered the LIBOR rate (in US dollars, at various terms) instead of the US deposit rate.

[33] We did not find a significant cointegration relationship between X_t, R_t^P and R_t^D or between X_t and R_t to treat (18) or (19) as an error correction model. There are some structural breaks in our 20 year data span that may explain this failure. Consistently with this, the *levels* of the returns did not appear to have enough explanatory power in the equations of Table 2.

Table 2. Augmented equations

	A_1	A_2	A_3	$A_3 - A_1A_2$	$H_0 : \boldsymbol{\beta} = 0$	p_R	$\bar{R}^2$
Mexico (1985.Q4 to 2005.Q3, $N = 77$)							
ex-ante	0.291*	0.202*	−0.273*	−0.331*	11.50*	2	0.554
	(0.096)	(0.066)	(0.092)	(0.113)	[0.000]		
ex-post	0.129	0.287*	−0.240*	−0.278*	22.56*	2	0.565
	(0.108)	(0.089)	(0.080)	(0.091)	[0.000]		
Peru (1980.Q1 to 2005.Q3, $N = 89$)							
ex-ante	0.242*	0.195*	0.003	−0.047*	9.086*	3	0.435
	(0.043)	(0.047)	(0.053)	(0.015)	[0.000]		
ex-post	0.501*	0.138**	−0.027	−0.069**	30.85*	2	0.649
	(0.098)	(0.083)	(0.068)	(0.036)	[0.000]		
Poland (1985.Q4 to 2005.Q3, $N = 68$)							
ex-ante	0.449*	0.132*	−0.002	−0.059*	1.638	3	0.275
	(0.043)	(0.058)	(0.049)	(0.022)	[0.203]		
ex-post	0.586*	0.164*	−0.123	−0.096*	2.402**	4	0.394
	(0.077)	(0.070)	(0.160)	(0.043)	[0.099]		
Uruguay (1985.Q1 to 2005.Q3, $N = 80$)							
ex-ante	0.252*	0.280*	−0.109	−0.070**	3.153*	3	0.124
	(0.104)	(0.114)	(0.140)	(0.038)	[0.049]		
ex-post	0.267*	0.349*	−0.073	−0.093**	2.189	2	0.152
	(0.103)	(0.117)	(0.143)	(0.047)	[0.119]		

2SLS estimates. Instruments for R_t^D and R_t^P (and for the *ex-ante* R_{t-1}^D and R_{t-1}^P) are oil prices changes, US GDP growth and lagged values of these and the R-variables. Figures in parentheses are robust (consistent) standard errors. * [**] denotes significance at a 5% [10%] level. Figures in the $H_0 : \boldsymbol{\beta} = 0$ column are F-statistics, p-values shown in braces. Lag length p_R was chosen to minimize the Schwarz criterion. For Peru, Poland and Uruguay, we set $A_3 = 0$ to compute the third central moment and its standard deviation. Diagnostic tests suggested well-behaved residuals. All regressions include a constant.

markets have allowed a very heterogenous group of agents – from a large firm that uses state-of-art portfolio management techniques to an uninformed individual, a blacksmith, who bases their portfolio decisions simply on their own experience and limited information – to participate in the same market. This paper shows that such an heterogeneity turns out to be enough to generate persistence in dollarization ratios upon aggregation. Empirical evidence from three Latin American countries and Poland supports this claim.

The presence of heterogeneity in individual dollarization decisions has interesting policy implications. Ize and Levy Yeyati (2003) conclude sensibly that a necessary and sufficient condition for dedollarization is higher exchange rate flexibility. In our setup this condition is not sufficient (though we reckon it is necessary), as there may exist a mass of individuals that do not respond at all to such a volatility. This makes a case for a more active policy on improving the communication skills of the central bank, in order to better convey its policy of more flexible exchange rates and possibly its commitment to price stability to a broader set of agents, specially to those regarded as uninformed. In this way the policymaker would be contributing to reduce individual heterogeneity and thus aggregate persistence.

This policy implication is particularly relevant for developing economies with an inflation targeting regime or for those evaluating moving towards this regime, as it heavily relies upon transparency and communication strategies. Our analysis suggests that the benefits of the such a policy regime in reducing dollarization may be condemned to be limited, unless the central bank effectively communicates the implications and benefits of such a regime to the less informed segment of participants in the dollar market.

References

Agénor, Pierre-Richard and Nadeem U. Haque, 1996, "Macroeconomic Management with Informal Financial Markets", *International Journal of Finance and Economics*, 1(2), 87-101.

Aoki, Kosuke, 2003, "On the Optimal Monetary Policy Response to Noisy Indicators", *Journal of Monetary Economics*, 50(3), 501-523.

Baillie, Richard, 1996, "Long Memory Processes and Fractional Integration in Econometrics", *Journal of Econometrics*, 73(1), 5-59.

Calvo, Guillermo A. and Carmen M. Reinhart, 2002, "Fear of Floating", *Quarterly Journal of Economics*, 107(2), 379-408.

Céspedes, Luis F., Roberto Chang and Andrés Velasco, 2004, "Balance Sheets and Exchange Rate Policy", *American Economic Review*, 94(4), 1183-1193.

Cook, David, 2004, "Monetary Policy in Emerging Markets: Can Liability Dollarization Explain Contractionary Devaluations?", *Journal of Monetary Economics*, 51(6), 1155-1181.

De Nicoló, Gianni, Patrick Honohan and Alain Ize, 2005, "Dollarization of Bank Deposits: Causes and Consequences", *Journal of Banking and Finance*, 29(7), 1697-1727.

Evans, George W., and Seppo Honkapohja, 2001, *Learning and Expectations in Macroeconomics*, Princeton University Press.

Granger, Clive W. J., 1980, "Long Memory Relationships and the Aggregation of Dynamic Models", *Journal of Econometrics*, 14(2), 227-238.

Guidotti, Pablo E. and Carlos A. Rodríguez, 1992, "Dollarization in Latin America: Gresham's Law in Reverse?", *IMF Staff Papers*, 39(3), 518-544.

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

Harvey, Andrew C. and Giuliano De Rossi, 2006, “Signal extraction”, ch. 31 in Mills, Terence C. and Kerry Patterson (eds.), *Palgrave Handbook of Econometrics*, Palgrave Macmillan.

Ize, Alain and Eduardo Levy Yeyati, 2003, “Financial Dollarization”, *Journal of International Economics*, 59(2), 323-347.

Kamin, Steven B. and Neil R. Ericsson, 2003, “Dollarization in Post-Hyperinflationary Argentina”, *Journal of International Money and Finance*, 22(2), 185-211.

Levy Yeyati, Eduardo, 2006, “Financial Dollarization: Evaluating the Consequences”, *Economic Policy*, forthcoming.

Lewbel, Arthur, 1994, “Aggregation and Simple Dynamics”, *American Economic Review*, 84(4), 905-918.

Ljungqvist, Lars and Thomas J. Sargent, 2000, *Recursive Macroeconomic Theory*, MIT Press.

Michelacci, Claudio, 2004, “Cross-sectional Heterogeneity and the Persistence of Aggregate Fluctuations”, *Journal of Monetary Economics*, 51(7), 1321-1352.

Morón, Eduardo and Diego Winkelried, 2005, “Monetary Policy Rules for Financially Vulnerable Economies”, *Journal of Development Economics*, 76(1), 23-51.

Pesaran, M. Hashem, 2003, “Aggregation of Linear Dynamic Models: An Application to Life-Cycle Consumption Models under Habit Formation”, *Economic Modelling*, 20(2), 383-415.

Quispe, Zenón, 2000, “Monetary Policy in a Dollarised Economy: the Case of Peru”, in Mahadeva, Lavan and Gabriel Sterne (eds.), *Monetary Policy Frameworks in a Global Context*, Routledge & Bank of England, London-New York.

Reinhart, Carmen M., Kenneth S. Rogoff and Miguel A. Savastano, 2003, “Addicted to Dollars”, NBER Working Paper No. 10015.

Savastano, Miguel A., 1996, “Dollarization in Latin America. Recent Evidence and some Policy Issues”, in Paul Mizen and Eric J. Pentecost (eds.), *The Macroeconomics of International Currencies: Theory, Policy, Evidence*, NH: Edward Elgar.

Sturzenegger, Federico, 1997, “Understanding the Welfare Implications of Currency Substitution”, *Journal of Economic Dynamics and Control*, 21(2/3), 391-416.

Uribe, Martín, 1997, “Hysteresis in a Simple Model of Currency Substitution”, *Journal of Monetary Economics*, 40(1), 185-202.

Zaffaroni, Paolo, 2004, “Contemporaneous Aggregation of Linear Dynamic Models in Large Economies”, *Journal of Econometrics*, 120(1), 75-102.

A The distribution of endowments and abilities

Our results were derived under the assumption that agents are homogenous in their endowments. In particular, we restricted the analysis to the case where each agent has an endowment of size one. Here, we show that our results hold for a more general case, one in which agents have different size of endowments, but where the distribution of abilities (a) across agents is correlated with that of the endowments. We regard this correlation as plausible in reality.

Consider equation (9). For the sake of argument, set $\mu = 0$ and let us assume that aggregate income is equal to one and that there are two agents in the economy: one with ability a_1 and income n_1 and the other with ability a_2 and income $n_2 = 1 - n_1$. Then,

$$(1 - a_i L)x_{it} = a_i \left(\frac{1}{\mathrm{v}_i + \sigma_\epsilon^2} \right) \xi_{it} \qquad \text{for } i = 1, 2 \tag{A1}$$

After generating a common lag polynomial for both process we have that

$$\begin{aligned} (1 - a_2 L)(1 - a_1 L)x_{1t} &= a_1 \left(\frac{1 - a_2 L}{\mathrm{v}_1 + \sigma_\epsilon^2} \right) \xi_{1t} \\ (1 - a_1 L)(1 - a_2 L)x_{2t} &= a_2 \left(\frac{1 - a_1 L}{\mathrm{v}_2 + \sigma_\epsilon^2} \right) \xi_{2t} \end{aligned} \tag{A2}$$

The aggregate level of dollar deposits, which coincides with the aggregate dollarization ratio, is $X_t = n_1 x_{1t} + n_2 x_{2t}$. Aggregate the equations in (A2) to get

$$(1 - a_1 L)(1 - a_2 L)X_t = n_1 a_1 \left(\frac{1 - a_2 L}{\mathrm{v}_1 + \sigma_\epsilon^2} \right) \xi_{1t} + n_2 a_2 \left(\frac{1 - a_1 L}{\mathrm{v}_2 + \sigma_\epsilon^2} \right) \xi_{2t} \tag{A3}$$

Define $\tilde{\xi}_{it} = n_i a_i (\mathrm{v}_i + \sigma_\epsilon^2)^{-1} \xi_{it}$ for $i = 1, 2$. Then, (A3) boils down to

$$X_t = (a_1 + a_2)X_{t-1} + a_1 a_2 X_{t-2} + \tilde{\xi}_{1t} - a_2 \tilde{\xi}_{1t-1} + \tilde{\xi}_{2t} - a_1 \tilde{\xi}_{2t-1} \tag{A4}$$

We have that the aggregate dollarization ratio follows an ARMA(2,1) process. This simple example can be generalizad to the case of N AR(1) process (hence N ability or endowment levels); in such a case the aggregate dollarization ratio follows an ARMA(N^*, $N^* - 1$) process, where $N^* \leq N$. We can increase the number of agents involved by simply replicating the individual behavior for a given ability a an arbitrary number of times. Therefore, the aggregation results derived in Appendix B go through under the assumption that the distribution of endowments is correlated to that of the abilities to process information. When $N \to \infty$, we get the limiting case exposed in Appendix C. The derivations in this appendix applies to the alternative version of our model that led to equation (16).

B Aggregation

The derivations herein follow Lewbel (1994) closely. To alleviate the notation we drop the i subscript in this appendix.

B.1 Equations (10) and (11)

Consider equation (9),

$$x_t = ax_{t-1} + m + u_t \tag{B1}$$

where $m = \mu(1-a)\mathrm{v}^{-1}$ and $u_t = a(\mathrm{v} + \sigma_\epsilon^2)^{-1}\xi_t$ depend on a. Since by construction ξ_t is a sequence of serially uncorrelated shocks, so is u_t. However, u_t is correlated across individuals.

Let $\mathcal{E}$ be the expectation operator across individuals, $\mathcal{E}[z] = \int z\,\mathrm{d}F(a)$, such that $X_t = \mathcal{E}[x_t]$, $M = \mathcal{E}[m]$ and $U_t = \mathcal{E}[u_t]$. Aggregation of (B1) renders

$$X_t = \mathcal{E}[ax_{t-1}] + M + U_t \tag{B2}$$

Define a random variable α_s, a scalar $A_s = \mathcal{E}[\alpha_s]$ and a recursion $\alpha_{s+1} = (\alpha_s - A_s)a$ with initial condition $\alpha_1 = a$. Note that for $s > 1$ the above recursion implies that $\alpha_s = a^s - \sum_{r=1}^{s-1} a^{s-r} A_r$. After taking $\mathcal{E}$ expectations we get equation (11) in the main text, where $m_s = \mathcal{E}[a^s]$ is the s-th moment of the distribution of a. Note also that

$$\begin{aligned} \mathcal{E}[\alpha_s x_{t-s}] &= A_s X_{t-s} + \mathcal{E}[(\alpha_s - A_s)x_{t-s}] \\ &= A_s X_{t-s} + \mathcal{E}[(\alpha_s - A_s)ax_{t-(s+1)}] + \mathcal{E}[(\alpha_s - A_s)m] + \mathcal{E}[(\alpha_s - A_s)u_{t-s}] \\ &= A_s X_{t-s} + \mathcal{E}[\alpha_{s+1}x_{t-(s+1)}] + \mathrm{cov}(\alpha_s, m) + \mathrm{cov}(\alpha_s, u_{t-s}) \end{aligned} \tag{B3}$$

where $\mathrm{cov}(\alpha_s, m)$ is the cross-sectional covariance of α_s and m which is time-invariant. On the other side, $\mathrm{cov}(\alpha_s, u_{t-s})$ is the cross-sectional covariance of α_s and u_{t-s} which is time dependent, but as this dependency comes from ξ_t, it is serially uncorrelated.

Equation (B3) shows a recursion between $\mathcal{E}[\alpha_s x_{t-s}]$ and $\mathcal{E}[\alpha_{s+1}x_{t-(s+1)}]$. After solving it,

$$\mathcal{E}[ax_{t-1}] = \sum_{r=1}^{\infty} A_r X_{t-r} + \sum_{r=1}^{\infty} \mathrm{cov}(\alpha_r, m) + \sum_{r=1}^{\infty} \mathrm{cov}(\alpha_r, u_{t-r}) \tag{B4}$$

Let $V_t = \sum_{r=1}^{\infty} \mathrm{cov}(\alpha_r, u_{t-r})$ and $\tilde{V} = E[V_t]$, where E is the expectation operator over time. Define also $\tilde{M} = M + \sum_{r=1}^{\infty} \mathrm{cov}(\alpha_r, m) + \tilde{V}$ and $\tilde{U}_t = U_t + V_t - \tilde{V}$. Then, after plugging (B4) into (B2) we get equation (10) in the main text, $X_t = \sum_{r=1}^{\infty} A_r X_{t-r} + \tilde{M} + \tilde{U}_t$, where $\tilde{U}_t$ is serially uncorrelated.[34] The underlying assumptions behind the aggregate equation (10) are thus, that $\tilde{M}$ and V_t are both finite or the sequences $\{\text{cov}(\alpha_s, m)\}_{s=1}^{\infty}$ and $\{\text{cov}(\alpha_s, u_{t-s})\}_{s=1}^{\infty}$ are absolute summable.

B.2 Equation (12)

Consider now equation (B1) in first differences

$$\Delta x_t = a\Delta x_{t-1} + u_t - u_{t-1} \tag{B5}$$

so that after aggregation, $\Delta X_t = \mathcal{E}[a\Delta x_{t-1}] + U_t - U_{t-1}$. Following the same procedure leading to equation (B4),

$$\mathcal{E}[a\Delta x_{t-1}] = \sum_{r=1}^{\infty} A_r \Delta X_{t-r} + V_t - V_{t-1} \tag{B6}$$

so that ΔX_t can be written as

$$\Delta X_t = \sum_{r=1}^{\infty} A_r \Delta X_{t-r} + (U_t + V_t) - (U_{t-1} + V_{t-1}) = \sum_{r=1}^{\infty} A_r \Delta X_{t-r} + U_t^{\dagger} \tag{B7}$$

which corresponds to the first-difference version of (10). The new aggregate error $U_t^{\dagger}$ is serially correlated and the coefficients are the same as those in (10).

B.3 Equation (17)

All the results derived above go through in aggregating (16). Note that this equation can be written as $x_t = ax_{t-1} + bS_t$ where a and b are individual specific coefficients whereas S_t is an aggregate figure. Thus, $\mathcal{E}[bS_t] = \mathcal{E}[b]S_t$, so $X_t = \mathcal{E}[ax_{t-1}] + \mathcal{E}[b]S_t$. Equation (B4) is now

$$\mathcal{E}[ax_{t-1}] = \sum_{r=1}^{\infty} A_r X_{t-r} + \sum_{r=1}^{\infty} \text{cov}(\alpha_r, b) S_t \tag{B8}$$

Call $\beta_0 = \mathcal{E}[b]$, $\beta_r = \text{cov}(\alpha_r, b)$, $\hat{U}_t = \sum_{r=0}^{\infty} \beta_r \varepsilon_{t-r}$ and recall that $S_t = R_t + \varepsilon_t$. Further mechanical manipulation leads to (17). The aggregate disturbance $\hat{U}_t$ is serially correlated.

[34] Pesaran (2003) shows that it is heteroscedastic, though.

C A brief note on fractional integration

Consider the univariate dynamic model

$$\Phi(L)(1-L)^d X_t = \Theta(L)\eta_t \tag{C1}$$

where L is the lag operator, $\eta_t \sim iid(0, \sigma_\eta^2)$ and d is the *differencing parameter*. When $d = 0$, X_t is stationary and follows an ARMA process, $\Phi(L)X_t = \Theta(L)\eta_t$. When $d = 1$, X_t has a unit root and hence follows an ARIMA process, $\Phi(L)\Delta X_t = \Theta(L)\eta_t$. More generally, when d takes non-integer values, X_t is said to be a fractionally integrated ARMA (ARFIMA) process. When $d \in (0, 0.5]$, the autocovariance function of X_t declines hyperbolically to zero, making X_t a stationary long-memory process. For $d > 0.5$, X_t is non-stationary (has infinite variance).

Granger (1980) has shown that under particular assumptions about $F(a)$ – the distribution of individual autoregressive coefficients – the aggregation of AR(1) processes like (9) leads to (C1).[35] In our empirical application, we simply imposed $d = 1$ and proceeded. If $d < 1$ truly, then we would have over-differentiated the data, with possible negative effects in our statistical inference.

Table B1 displays estimates of d and tests $H_0 : d = 0$ and $H_0 : d = 1$. We did not find enough evidence to reject $H_0 : d = 1$ whereas $H_0 : d = 0$ is systematically rejected.

Table B1. Estimated fractional integration parameter in dollarization ratios

		$H_0 : d = 0$		$H_0 : d = 1$	
	$\hat{d}$	t-stat	p-value	t-stat	p-value
Mexico	0.825	2.376	0.0491	0.505	0.6294
Peru	0.932	3.883	0.0037	0.282	0.7843
Poland	0.955	4.605	0.0025	0.219	0.8333
Uruguay	0.788	2.485	0.0378	0.667	0.5236

The estimation method is that of Geweke and Porter-Hudak (known as GPH). The asymptotic standard error of $\hat{d}$ is $\pi^2/6$ which is used to compute the t-statistics and p-values. Both tests ($H_0 : d = 0$ and $H_0 : d = 1$) are two-tailed. See Baillie (1996) for a review of ARFIMA modelling and for critics to the GPH estimator.

[35] See also Baillie (1996) and Zaffaroni (2004).