



**CENTRO DE
ESTUDIOS
MONETARIOS
LATINOAMERICANOS**

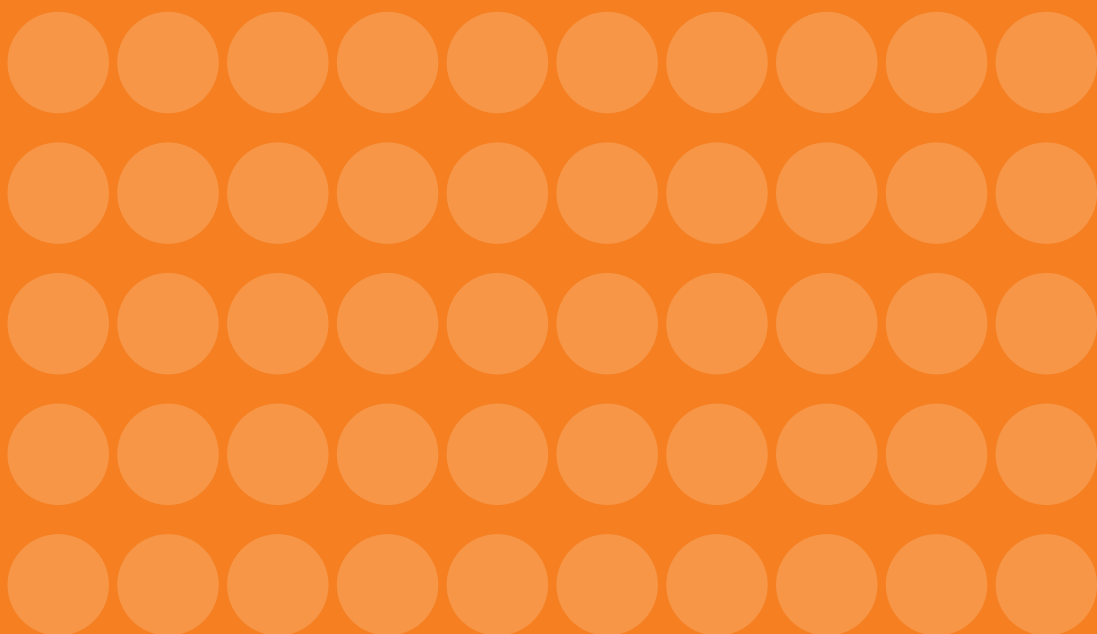
Asociación Regional de Bancos Centrales

ESTUDIOS

**¿Por qué la dolarización
es tan persistente?**

Paul Castillo y Diego Winkelried

Premio de Banca Central Rodrigo Gómez 2006



¿POR QUÉ LA DOLARIZACIÓN ES TAN PERSISTENTE?

Paul Castillo y Diego Winkelried

¿Por qué la dolarización es tan persistente?

PREMIO DE BANCA CENTRAL “RODRIGO GÓMEZ, 2006”

CENTRO DE ESTUDIOS MONETARIOS LATINOAMERICANOS
México, D. F. 2009

Primera edición, 2009

© Centro de Estudios Monetarios Latinoamericanos, 2009
Derechos reservados conforme a la ley
Durango n^o 54, México, D. F., 06700
ISBN: 978-607-7734-03-1

Impreso y hecho en México
Printed and made in Mexico

Presentación

El CEMLA se complace en publicar en español e inglés el trabajo que obtuvo el Premio de Banca Central Rodrigo Gómez, 2006, elaborado por los economistas peruanos Diego Winkelried y Paul Castillo.

Los autores inician su estudio con una revisión del estado de dolarización en economías latinoamericanas para valorar su persistencia. Esto último, es una de las características más importantes de la dolarización financiera y la que causa mayor preocupación a los bancos centrales. Aun después de haber alcanzado niveles de inflación bajos y estables los coeficientes de dolarización a menudo permanecen elevados en muchas economías. En este trabajo se muestra que la persistencia de la dolarización está relacionada con la heterogeneidad de los participantes en el mercado de depósitos en dólares, quienes poseen capacidades diferentes para procesar información. La relevancia empírica de este hecho se corrobora con datos provenientes de tres países de América Latina y de Polonia.

El CEMLA espera contribuir al conocimiento y profundización del fenómeno de la dolarización en la región con la difusión en español e inglés del estudio ganador del Premio de Banca Central Rodrigo Gómez 2006.

Paul Castillo labora en el Banco Central de Reserva del Perú (BCRP), donde desempeña el cargo de Subgerente de Diseño de Política Monetaria. Ingresó al BCRP en 1996, y desde esa fecha ha desempeñado diversos cargos. Cursó la maestría y el doctorado en Economía en la London School of Economics and Political Science, Inglaterra, y antes estudió Ingeniería Económica en la Universidad Nacional de Ingeniería, Perú. Ha sido profesor en diversas universidades. Actualmente enseña en la Maestría en Economía en la Universidad del Pacífico, en el Programa de Ingeniería Económica de la Universidad Nacional de Ingeniería, y en el Programa de Economía de la Universidad Privada de Ciencias Aplicadas. Asimismo, ha participado en diversos cursos, seminarios y conferencias dentro y fuera de Perú, donde también ha presentado investigaciones y publicaciones en sus áreas de interés (macroeconomía, política monetaria y economía aplicada). Diego Winkelried es MPhil in Economics (distinction) por la Universidad de Cambridge, Inglaterra, y Bachiller de Economía por la Universidad del Pacífico, Perú. Actualmente estudia el doctorado en la Universidad de Cambridge. Ha sido profesor de Matemáticas para Economistas en la Universidad del Pacífico y Analista del Departamento de Modelos Macroeconómicos del BCRP. Se le han otorgado premios y menciones en concursos de investigación internacionales (CEMLA, INTAL, Banco de Guatemala) y sus investigaciones han sido publicadas en diversos medios académicos. Es becario del Gates Cambridge Trust y *scholar* de St John's College, Cambridge. Las opiniones expresadas en este trabajo son responsabilidad de los autores y no representan la opinión de sus respectivas instituciones ni del CEMLA. El artículo original fue escrito en inglés, los errores derivados de la traducción son responsabilidad del CEMLA. Autor de contacto: D. Winkelried, <dw295@cam.ac.uk>.

1. Motivación

Si bien, es cierto que la dolarización es un área de investigación relativamente nueva, la experiencia de muchas economías latinoamericanas y en transición durante la década de los noventa han inspirado una amplia y creciente literatura sobre el tema.¹ La dolarización está normalmente asociada a la sustitución parcial de la moneda nacional por una moneda extranjera (el dólar estadounidense) como depósito de valor, lo cual difiere de la sustitución monetaria que se refiere al uso de la moneda extranjera como medio de cambio.

En este documento, dolarización significa *dolarización de depósitos*,² que lleva en última instancia a la dolarización de créditos y de allí a la vulnerabilidad del sistema financiero de países cuyas economías están altamente dolarizadas. Como señalan Cook (2004) y Céspedes *et al.* (2004), la eficacia de la política monetaria en economías pequeñas y abiertas con tipos de cambio flexibles se compromete por los efectos negativos de la hoja de balance generados por la dolarización. En este caso, las depreciaciones reales repentinas pueden resultar perjudiciales para la actividad económica al reducir el patrimonio neto de las empresas y generar efectos adversos en la inversión. Este hecho explica en parte el *temor a flotar* de los bancos centrales (Calvo y Reinhart, 2002; Morón y Winkelreid, 2005).

Una de las características más sobresalientes de la dolarización y quizás aquella que causa mayor preocupación a los formuladores de política económica es su *persistencia*. Se ha documentado que la dolarización aumenta de forma brusca durante episodios de excesiva inestabilidad macroeconómica y que permanece elevada de manera persistente, incluso después de que se alcanza con éxito la estabilidad.³ Una primera explicación de la histéresis es la falta de confianza en los activos

¹ Ver De Nicoló *et al.* (2005), y Levy Yeyati (2006), y las referencias bibliográficas detalladas en este documento.

² Conocida también como sustitución de activos (Reinhart *et al.* 2003) o dolarización financiera (Ize y Levy Yeyati, 2003).

³ Ver Guidotti y Rodríguez (1992), Savastano (1996), Quispe (2000) y Karmin y Ericsson (2003).

denominados en moneda nacional debido a los traumas generados por inflaciones, devaluaciones, y crisis bancarias del pasado. Sin embargo, esta explicación no resulta muy consistente debido a los sólidos fundamentos macroeconómicos observados en economías muy dolarizadas (en particular, el Perú y algunas economías en transición a inicios de los años 2000).

Una estrategia alternativa para explicar este dilema es modificar los modelos de sustitución monetaria existentes basados en costos de ajuste o externalidades de red. Guidotti y Rodríguez (1992), Sturzenegger (1997) y Uribe (1997) desarrollan modelos en los que el costo de usar el dólar para transacciones depende negativamente del coeficiente de sustitución monetaria, así una vez que las transacciones están dolarizadas, ya no hay beneficio en volver a usar la moneda nacional si otros siguen usando el dólar. Una limitación obvia de este enfoque es que éste se refiere a la función de medio de cambio de la moneda y no a la función de depósito de valor. Además, estos modelos dependen en gran medida de una variable *rachet*, un acervo de conocimientos que conduce a la persistencia; así aunque estos explican nítidamente la tendencia creciente en el coeficiente de dolarización, no parecen útiles para aconsejar a los formuladores de política acerca de cómo desdolarizarse, debido a que implican una reducción poco probable en el acervo de conocimiento.

Ize y Levy Yeyati (2003) ofrecen un marco diferente para modelar la dolarización. Ellos utilizan un enfoque de selección de cartera y derivan una cartera de mínima varianza (minimum variance portfolio, MVP) que depende de la volatilidad relativa de la inflación y de la depreciación del tipo de cambio real. La dolarización persistirá aun cuando la inflación esté baja y estable en la medida que la volatilidad de la depreciación del tipo de cambio real sea menor que la de la inflación. Sin embargo, este modelo es estático mientras que la persistencia es inherentemente un fenómeno dinámico. A nuestro parecer, el enfoque MVP –el cual es bastante popular y ha sido probado con éxito para explicar la variación transversal de la dolarización⁴ no ha sido diseñado para tratar

⁴ Ize y Levy Yeyati (2003) proporcionan evidencia empírica de que la MVP tiene algún poder para explicar el nivel promedio de dolarización

con la dinámica, ya que la MVP, el nivel de equilibrio subyacente de dolarización, depende de los momentos *incondicionales*.⁵

Curiosamente un hecho que al parecer los investigadores han pasado por alto es la naturaleza misma de los participantes del mercado de depósitos en dólares en economías dolarizadas: los depositantes en dólares son muy heterogéneos; van desde grandes a pequeñas empresas, a organizaciones sin fines de lucro e incluso a individuos (con distinto nivel de ingresos).⁶ Los costos de participación son casi nulos debido a la liberación y desregulación, y de manera sobresaliente a la aparición de intermediarios de cambios informales, conocidos como *cambistas* en muchos países de América Latina, quienes se benefician de la compra y venta de dólares con márgenes más pequeños que los del sistema bancario.⁷ Un típico cambista porta una cantidad limitada de dinero para sus negocios (usualmente, entre 3.000 y 5.000 dólares estadounidenses) ya que tiene como objetivo enfrentar la demanda de dólares para individuos o pequeñas empresas que no desean pagar las primas más altas que cobran los bancos para dolarizar sus ahorros.⁸ En resumen, la participación en el mercado de cambios se vuelve independiente de la escala de la transacción y por ello se masifica.

El propósito de este documento es mostrar el hecho de que la heterogeneidad de los depositantes puede explicar de manera simple la persistencia de la dolarización financiera. Como señala Granger (1980), las diferencias en las dinámicas individuales conducen a una persistencia agregada. De ese

entre países. De Nicolás *et al.* (2005) extiende este análisis empírico considerado una muestra más grande de países.

⁵ La histéresis de la dolarización se observa en varios países con alta volatilidad del tipo de cambio real, por ejemplo, Rusia. La razón de esta aparente contradicción con el enfoque de cartera puede ser que resulta muy difícil obtener estimaciones seguras de las varianzas incondicionales que componen la MVP.

⁶ Una excepción es Sturzenegger (1997), donde se estudian las implicaciones de desigualdad de ingreso en la sustitución monetaria, sin hacer referencia a la dolarización de depósitos.

⁷ Agénor y Haque (1996) proporcionan una descripción del mercado de cambios informal.

⁸ Incluso para las grandes empresas, les puede resultar rentable comerciar con un grupo de (bien organizados) cambistas.

modo, en tanto es razonable esperar que la dinámica de una cartera óptima de monedas de un experto financiero difiera de la de un no experto, un coeficiente de dolarización agregada persistente surge naturalmente. Desde luego, existen varias diferencias entre un experto financiero y un no experto, sin embargo, si ambos acceden al mercado de cambios casi de manera gratuita, la diferencia relevante para nuestro análisis se centra en su capacidad para procesar información, y por lo tanto, para tomar decisiones de ahorro informadas.⁹

Ofrecemos una extensión simple del enfoque MVP al considerar que los depositantes son heterogéneos en su capacidad para procesar información y para predecir rendimientos reales. En particular, se asume que los depositantes reciben información *ruidosa* y que la cantidad de ruido que reciben es idiosincrásica. Resulta que según tales supuestos, el coeficiente de dolarización óptimo de un individuo sigue un proceso $AR(1)$, donde el grado de persistencia está asociado con la habilidad del depositante para extraer información. En particular, los coeficientes de dolarización individuales serán menos (más) persistentes y responderán más (menos) en impacto a cambios en las señales, cuando la habilidad del depositante para procesar información sea mayor (menor). Mediante la agregación de las dinámicas individuales de los depositantes, mostramos que la dolarización financiera puede ser muy persistente cuando la heterogeneidad entre los depositantes es alta, aunque a nivel individual la persistencia sea moderada.

La estructura del documento es la siguiente. En la sección 2, se analizan brevemente estos temas usando datos de Perú y de Polonia.¹⁰ Por razones que se explican más adelante, estos casos calzan muy bien con nuestro propósito de ilustrar nuestra afirmación acerca de la interacción entre heterogeneidad

⁹ Las diferencias de ingreso también pueden ser importantes si la brecha de ingreso entre el experto financiero y el no experto es amplia. Sin embargo, encontramos que en economías dolarizadas, la participación en el mercado de depósitos en dólares de muchas empresas y varios individuos es similar.

¹⁰ Las cifras usadas en la sección 2, fueron proporcionadas por el Banco Central de Reserva del Perú y el Banco Nacional de Polonia. Los hechos tratados están registrados en los informes anuales de estas instituciones.

individual y persistencia agregada. Además, nos dan una idea de cómo participan varios tipos de depositantes en el mercado de depósitos en dólares.

En la sección 3 se desarrolla un modelo estilizado donde los depositantes enfrentan información ruidosa y difieren en su capacidad para pronosticar cuando se toman decisiones de cartera. Un resultado importante de este mecanismo es que la dinámica de las carteras óptimas de los individuos depende de sus errores de predicción acerca de los rendimientos futuros en dólares. Resulta óptimo que los agentes sean prudentes cuando modifican la composición de sus depósitos por monedas debido a que existe incertidumbre en la calidad de los datos que éstos reciben. Esta actitud de precaución se refleja en cambios suaves en la composición de sus carteras. Finalmente mostramos como a través de la agregación de las decisiones de dolarización individual, es posible generar un coeficiente de dolarización para toda la economía bastante persistente.¹¹

En la sección 4 se evalúan las hipótesis empíricas del modelo teórico y se encuentra evidencia favorable a partir de los datos agregados de tres países latinoamericanos y de Polonia. En particular, los datos sugieren que las distribuciones de las *capacidades de predicción* detrás de los coeficientes de dolarización agregada son muy dispersas. Consideramos que este resultado es coherente con la idea de que los expertos financieros comparten el mercado de depósitos en dólares con no expertos que ahorran en dólares. En la sección 5 se reformula el modelo teórico para reflejar una fuente distinta de heterogeneidad de información, y proporcionar evidencia empírica adicional.

La sección 6 concluye y ofrece algunas recomendaciones de política. Las derivaciones y resultados complementarios se muestran en el apéndice.

¹¹ Nuestro enfoque está relacionado con otra corriente de la literatura. Por ejemplo, Lewbel (1994) utiliza información agregada para evaluar la heterogeneidad en las dinámicas de consumo, en tanto Michelacci (2004) explica el alto grado de persistencia del producto con heterogeneidad transversal de empresas productivas.

2. Dos casos ilustrativos

Como documenta Savastano (1996), la dolarización emerge progresivamente en respuesta a la inestabilidad macroeconómica, en particular a los altos niveles de inflación, y muestra un patrón bien definido: los primeros agentes que se dolarizan reemplazan la moneda nacional en su función de depósito de valor, ahorrando dólares por lo general fuera del sistema financiero (*debajo del colchón*). Luego, el dólar se usa en algunas transacciones, típicamente aquellas asociadas con bienes raíces y bienes durables, y con el tiempo algunos precios llegan a fijarse en dólares. Posteriormente la mayoría de los gobiernos permiten a los bancos que acepten depósitos en moneda extranjera para evitar la desintermediación financiera.¹² La experiencia real de varios países muestra que en un año la dolarización puede crecer bastante rápido. Ver gráficas I(A) y II(A).

Por otro lado, los episodios de desdolarización (es decir, la reducción sostenida del coeficiente de dolarización) no son muy comunes, y por eso no existe ningún patrón bien establecido en la literatura. Sin embargo, si alguna vez sucede, es probable que la desdolarización sea un proceso lento. El análisis de estos eventos, a diferencia de aquellos donde crece la dolarización, brinda información muy valiosa acerca de como depositantes distintos deciden acerca de la composición por monedas de sus ahorros y también sobre cómo responden a los cambios en el entorno macroeconómico.

2.1. Perú a inicios de la década del 2000

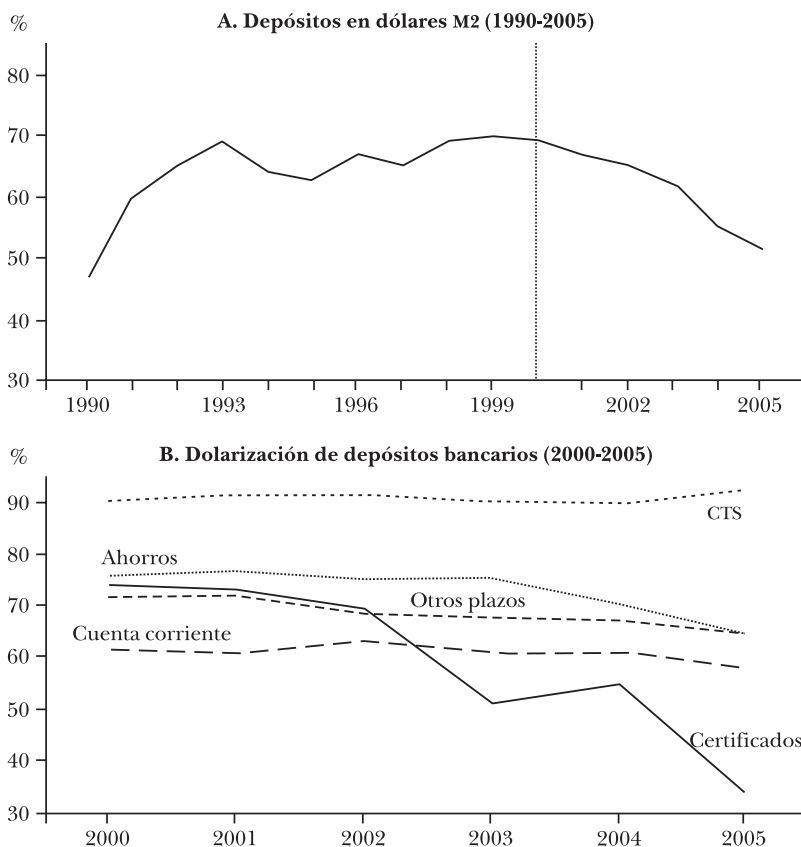
Aunque la experiencia de dolarización en el Perú presenta varias de las características antes mencionadas, esta tiene sus propias peculiaridades.¹³ Como se muestra en la gráfica I(A) en 1991 (después de una hiperinflación con cuatro dígitos en

¹² Ver también Karmin y Ericsson (2003), De Nicoló *et al.* (2005) y Levy Yeyati (2006).

¹³ Para una revisión histórica y detallada de la experiencia de dolarización en el Perú ver Quispe (2000).

1990) el coeficiente de dolarización era 60% y fluctuó alrededor de 65% y 70% por una década. Sin embargo, a partir del año 2000 se ha observado una reducción sostenida de la dolarización, hasta alcanzar niveles cercanos a 50% en el 2005. Es cierto que 50% es aún un porcentaje elevado, pero existen varios hechos interesantes que vale la pena destacar detrás de esta reciente reducción.

GRÁFICA I. DOLARIZACIÓN DE LOS DEPÓSITOS EN PERÚ, 1990-2005



FUENTE: Banco Central de Reserva del Perú.

Existen al menos dos factores que explican esta disminución. Primero, después de ocho años de anunciar las metas de inflación dentro de un régimen de metas monetarias (desde 1994) y después de cinco años de haber logrado niveles

de inflación de un solo dígito, el Banco Central de Reserva del Perú anunció la adopción por completo del esquema de metas explícitas de inflación en el 2002. Esta decisión ha contribuido a anclar las expectativas de inflación y ha reducido la volatilidad de la inflación y de la tasa de interés nominal. Segundo, entre el 2001 y el 2005 el tipo de cambio real y nominal se apreciaron (6.2% y 5.1%) como resultado de un entorno internacional favorable: fuerte incremento en los términos de intercambio que condujeron a un auge en las exportaciones y bajas tasas de interés internacional. En resumidas cuentas, el rendimiento real de los depósitos en dólares en relación con el rendimiento de los depósitos en moneda nacional se ha reducido en forma considerable desde inicios de la década del 2000.

La gráfica I(B) muestra la dolarización por tipo de depósito: vista, ahorros, y un desglose en los depósitos a plazo entre certificados de depósito y depósitos de Compensación por Tiempo de Servicios, CTS y otros. Una mirada a la gráfica revela que tanto los depósitos a la vista como los depósitos CTS no han reaccionado al cambio reciente en la tendencia de los rendimientos reales en dólares. Los depósitos a la vista representan cerca de 20% de los depósitos totales y son altamente líquidos, casi transaccionales, por lo que tienden a reaccionar menos a cambios en los rendimientos relativos. Por otro lado, los CTS son la versión peruana de un seguro de desempleo; por ley, sus tenedores son exclusivamente individuos y pueden ser utilizados sólo cuando el individuo está desempleado. Los depósitos CTS han reaccionado incluso menos que los depósitos a la vista a la reducción de los rendimientos en dólares, lo que resulta difícil de explicar.

La gráfica también muestra una tendencia descendente moderada en la participación de los depósitos de ahorro y a plazos en dólares. Cerca de 80% de los depósitos de ahorro y casi la mitad de los depósitos a plazo son mantenidos por individuos. Desde 2001 hasta 2005, ambos coeficientes han disminuido cerca de 10%. Lo que resalta en la gráfica I(B) es la fuerte reacción del coeficiente de los certificados de depósito que ha disminuido en casi 40% y que explica en gran medida la reducción del coeficiente de dolarización agregado en la gráfica I(A). El hecho interesante es que pese a que los

certificados de depósito tiene un plazo similar al de los CTS y al de los depósitos a plazo, los certificados de depósito pertenecen principalmente a empresas y no a individuos.

2.2. Polonia hacia una economía de mercado

La experiencia polaca es considerada como la experiencia más exitosa de transición desde una economía planificada hacia una economía de mercado, y es también un ejemplo sobresaliente de desdolarización. A finales de los años ochenta, Polonia estaba al borde de una profunda crisis económica. Las enormes distorsiones en los precios relativos y los déficits fiscales acumulados, heredados de los años de planificación central indujeron a un rápido aumento en la inflación que alcanzó un máximo histórico de 550% en 1989. En respuesta a este entorno macroeconómico inestable, los coeficientes de dolarización aumentaron rápidamente, de valores cercanos a 20% en 1985 hasta un pico de 60% en 1989. Esto se muestra en la gráfica II(A).

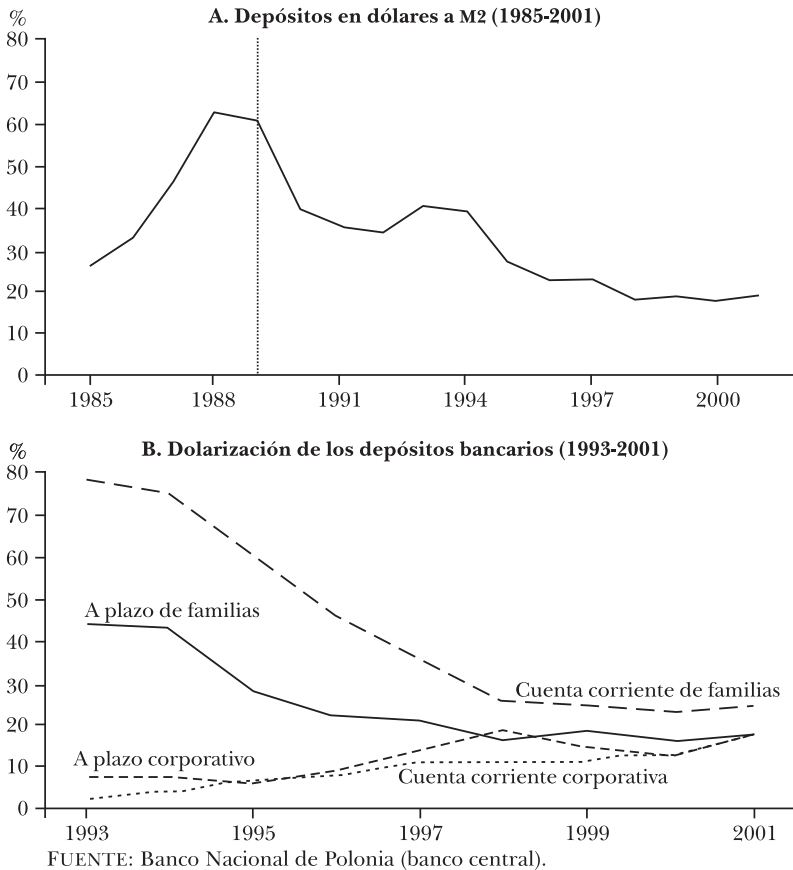
Después de introducir una serie de reformas a favor del mercado y de un programa de estabilización (denominado *terapia de choque*),¹⁴ los coeficientes de dolarización descendieron a promedios de 40% a fines de 1993, en conjunto con la reducción de la inflación (de 500% a 36%). A medida que las condiciones macroeconómicas mejoraron, se llevó a cabo una serie de reformas institucionales adicionales. En forma notable, en 1997 se concedió la independencia al Banco Central de Polonia así como un objetivo bien definido: garantizar la estabilidad de precios. La dolarización disminuyó aún más, hasta alcanzar un valor de 18% en 2001 comparable con los indicadores observados en economías de los países europeos desarrollados, como el Reino Unido.

Una característica común entre la experiencia polaca y la

¹⁴ Una drástica serie de reformas institucionales y de mercado fueron puestas en marcha en 1990: el gobierno liberalizó el control de casi todos los precios, eliminó la mayoría de los subsidios, abolió la designación administrativa de recursos a favor del comercio, promocionó el establecimiento libre de empresas privadas, liberalizó las relaciones económicas internacionales e introdujo la convertibilidad de moneda interna con una devaluación de la moneda de 32%.

peruana antes mencionada es la heterogeneidad que se observa en la dinámica de dolarización entre tipos de depósitos. La gráfica II(B) revela que a finales de 1993, la diferencia entre los coeficientes de dolarización entre las familias y las empresas fue 70% para depósitos a plazo y 40% para depósitos a la vista. Estas diferencias permanecieron en un rango de 20% por más de cuatro años.

GRÁFICA II. DOLARIZACIÓN DE LOS DEPÓSITOS EN POLONIA, 1985-2001



2.3. Moraleja

Las diferencias entre la forma como los individuos y las empresas deciden acerca de la composición de sus carteras son obvias. Usualmente las empresas tienen más recursos

asignados al manejo de sus fondos, mientras que los individuos a menudo basan sus decisiones en su experiencia, la de sus vecinos y el acceso limitado a la información. Incluso, es posible encontrar que el proceso de toma de decisiones es distinto dentro del grupo de individuos como dentro del de las empresas. La breve reseña de la experiencia de Polonia y Perú ilustra nuestro principal argumento de que tales diferencias explican la heterogeneidad en las decisiones de dolarización. A continuación analizamos como esto se traduce en la persistencia.

3. Un modelo simple

Usamos un marco conceptual simple para mostrar cómo la combinación de información imperfecta y ruidosa sobre los rendimientos reales de activos externos, y especialmente la heterogeneidad entre los participantes del mercado pueden generar un grado persistente de dolarización.

En el modelo, la economía está poblada por un gran número de individuos casi idénticos. Por simplicidad se asume que todos tienen la misma dotación, que se normaliza a 1, y las mismas preferencias, pero que difieren en su habilidad para procesar información y por lo tanto en sus expectativas sobre los resultados futuros.¹⁵

Cada periodo los agentes escogen la composición de su cartera entre dos activos, uno que ofrece un rendimiento real fijo R^p , que está denominado en moneda nacional (*peso* de ahora en adelante) y el otro denominado en dólares con rendimiento real R_t^D . Para simplificar la exposición del modelo nos enfocamos en el exceso real de rendimiento del dólar sobre el activo peso.

$$(1) \quad R_t = R_t^D - R^p.$$

3.1. Decisión de cartera

Los depositantes son adversos al riesgo. El individuo i destina una cantidad x_{it} de sus ahorros a adquirir el activo en dólares y el restante $1 - x_{it}$ para adquirir el activo en pesos. Según Ize y Levy Yeyati (2003), se supone una función de utilidad promedio-varianza simple. Puesto que la decisión de cartera es *ex ante* y está basada en información imperfecta sobre rendimientos reales, la utilidad para el individuo i está definida en términos de expectativa condicional para el periodo $t + 1$ con información hasta el periodo t .¹⁶

¹⁵ Nuestros análisis se mantiene aun para agentes con dotaciones heterogéneas –es decir, desigualdad riqueza/ingreso dado– que están correlacionadas con la capacidad para procesar información. Ver el apéndice A para detalles.

¹⁶ Hemos impuesto un valor de uno al parámetro de aversión al riesgo

$$\begin{aligned}
 (2) \quad U_{it} &= E_i \left[x_{it} R_{t+1}^D + (1-x_{it}) R^P \right] - \left(\frac{1}{2} \right) \text{var}_i \left[x_{it} R_{t+1}^D + (1-x_{it}) R^P \right] \\
 &= E_i \left[x_{it} R_{t+1} + R^P \right] - \left(\frac{1}{2} \right) \text{var}_i \left[x_{it} R_{t+1} \right] = x_{it} \hat{r}_{it+1} + R^P - \left(\frac{1}{2} \right) (x_{it})^2 v_{it+1}.
 \end{aligned}$$

Donde \hat{r}_{it+1} y v_{it+1} son el promedio y la varianza del exceso de rendimiento R_t que el individuo i espera para el periodo $t+1$, condicional a la información hasta el periodo t .

El valor de x_{it} que maximiza (2) es:

$$(3) \quad x_{it} = \frac{\hat{r}_{it+1}}{v_{it+1}}.$$

Así, los agentes aumentarían sus depósitos en dólares cuando esperan un rendimiento real más alto sobre este activo para la misma varianza esperada o cuando esperan una varianza más baja dado un nivel de exceso de rendimiento.

3.2. Pronósticos

En el periodo t , el exceso de rendimiento R_t no puede observarse de manera perfecta. Lo que se conoce públicamente es una versión de señal ruidosa de R_t , $S_t = R_t + \varepsilon_t$ donde $\varepsilon_t \sim iid(0, \sigma^2)$. En la práctica, estas variables pueden ser entendidas como diferenciales *ex ante* y *ex post* respectivamente. Como se muestra en la ecuación (3), las únicas piezas de información relevantes para la decisión de cartera son el exceso de rendimiento y su varianza. Para hacerlo más sencillo postulamos que cada agente se enfoca directamente en pronosticar R_t y no en sus componentes (R_t^D o R^P , que puede implicar el pronóstico de la inflación, la depreciación, el riesgo de confiscación, etc.).

Cada individuo tiene un *modelo para pronosticar* de la forma siguiente:

$$(4) \quad \begin{aligned}
 r_{it+1} &= \mu(1-a_i) + a_i r_{it} + \omega_{it+1} & \omega_{it} &\sim iid(0, \sigma_\omega^2) \\
 S_t &= r_{it} + \varepsilon_{it} & \varepsilon_{it} &\sim iid(0, \sigma_\varepsilon^2)
 \end{aligned}$$

Dado que S_t es un indicador ruidoso, el individuo i tiene que extraer primero la señal r_{it} (es decir, *pronóstico de muy*

en la función de utilidad. Este supuesto no afecta nuestros resultados.

corto plazo) y luego predecir su promedio y varianza para implementar (3).

Como se puede observar en (4), cada individuo recibe un valor de $|a_i| \leq 1$ obtenido de una distribución $F(a)$, para realizar sus predicciones, y este único valor determina el modelo de predicción completo. Esta es la única fuente de heterogeneidad (transversal) en este modelo. El parámetro μ es común para todos los individuos, lo que significa que ellos comparten el mismo pronóstico de largo plazo. También se asume que las varianzas de las perturbaciones, σ_w^2 y σ_ϵ^2 , son las mismas para todos los individuos.

Para ser más precisos, lo que viene a diferenciar a los individuos es la velocidad con que ajustan sus pronósticos a corto plazo a medida que reciben nueva información. Interpretamos esta heterogeneidad como las diferencias en la capacidad de las personas para realizar pronósticos,¹⁷ que se traducen de manera directa en diferencias de cartera entre los individuos.

Definimos $v_{it} = E\left[(r_{it} - \hat{r}_{it})^2\right]$ como el error cuadrático medio (ECM) de la variable predictiva \hat{r}_{it} . Los resultados estándar de la literatura de extracción de señales nos conducen a la siguiente regla óptima de predicción.¹⁸

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

donde el valor pronosticado de r_{it} para el siguiente periodo es la proyección del valor pronosticado de hoy más una corrección, una actualización de $\xi_{it} = \epsilon_{it} + (r_{it} - \hat{r}_{it})$ que es proporcional al último error de predicción incurrido. El valor de k_{it} está determinado por el coeficiente (ajustado) del ECM de \hat{r}_{it} a la varianza del indicador ruidoso.

¹⁷ Por ejemplo, si el verdadero proceso generador de datos de R_t es una caminata aleatoria, entonces las predicciones de los agentes con $a_i = 1$ superará a las predicciones de otros individuos (el verdadero proceso generador de datos de R_t , es irrelevante para la derivación de nuestros principales resultados).

¹⁸ El lector familiarizado con la literatura acerca de la modelación de espacio-tiempo notará que las recursiones (5) y (7) son aplicaciones directas del filtro de Kalman. Ver Ljungqvist y Sargent (2000) cap. 2 y de 21, y Harvey y De Rossi (2006) para detalles adicionales.

$$(6) \quad k_{it} = a_i \left(\frac{v_{it}}{v_{it} + \sigma_\epsilon^2} \right).$$

El ECM de \hat{r}_{it} evoluciona según la siguiente recurrencia:

$$(7) \quad v_{it+1} = \frac{v_{it}(a_i^2 \sigma_\epsilon^2 + \sigma_\omega^2) + \sigma_\epsilon^2 \sigma_\omega^2}{v_{it} + \sigma_\epsilon^2}.$$

De la ecuación (7) se tiene claro que $v_{it+1} = f(v_{it})$. Existe un punto fijo tal que $v_i = f(v_i)$ ¹⁹ y además, dado que $f'(v_i) < 1$, este es estable en forma global: independientemente de la condición inicial v_{i0} tenemos que $v_{it} \rightarrow v_i$ y en consecuencia $k_{it} \rightarrow k_i = a_i v_i (v_i + \sigma_\epsilon^2)^{-1}$ como $t \rightarrow \infty$. Esto significa que a medida que t es más grande, es decir, a medida que cada individuo ha realizado el ejercicio de extracción de señales varias veces, el proceso de actualización definido en las expresiones (5) y (7) converge a una regla de equilibrio.²⁰ Si asumimos que este proceso de aprendizaje comenzó mucho tiempo antes de un periodo t , entonces con seguridad v_{it} y k_{it} pueden considerarse como constantes. Este hecho simplifica los cálculos de manera considerable sin comprometer nuestras conclusiones.

Para entender mejor como la heterogeneidad entre agentes afecta sus pronósticos (y por tanto sus carteras), se asume por un momento que $\mu = 0$ y se resuelve (5) de forma recursiva para obtener.

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

Es claro de esta expresión de rezagos distribuidos geométricamente, que diferentes realizaciones de a están asociados con formas distintas de ponderar la información disponible (los indicadores ruidosos hasta un periodo t) para producir un pronóstico.²¹

¹⁹ El punto fijo es la raíz positiva de $v_i^2 + [(1 - a_i^2) \sigma_\epsilon^2 - \sigma_\omega^2] v_i - \sigma_\epsilon^2 \sigma_\omega^2 = 0$.

²⁰ La convergencia es monótonica ($v_{it} \geq v_{it+1} \geq v_i$) porque v_{it+1} se basa en más información que v_{it} . Este esquema de actualización es *E-estable*, en el sentido de Evans y Honkapohja (2001) en el aprendizaje adaptativo.

²¹ Como se precisa en Harvey (1989, cap. 4), el modelo de pronóstico popular converge al método de suavizamiento exponencial (exponential smoothing, ES) cuando $a \rightarrow 1$. Sin embargo, el esquema explicado aquí es

3.3. Dinámicas individuales

Utilizamos el hecho de que $v_{it} \rightarrow v_i$, $k_{it} \rightarrow k_i$ y la regla de actualización o pronóstico óptima (5) de tal forma que la inversión en dólares óptima (3) se reduce a:

$$(8) \quad x_{it} = \frac{\hat{r}_{it+1}}{v_{it+1}} = \frac{\hat{r}_{it+1}}{v_i} = \frac{\mu(1-a_i) + a_i \hat{r}_{it} + k_i \xi_{it}}{v_i} = a_i x_{it-1} + \frac{\mu(1-a_i)}{v_i} + \left(\frac{k_i}{v_i} \right) \xi_{it}.$$

Después de reemplazar (6) en (8), obtenemos:

$$(9) \quad x_{it} = a_i x_{it-1} + \frac{\mu(1-a_i)}{v_i} + a_i \left(\frac{1}{v_i + \sigma_\epsilon^2} \right) \xi_{it}.$$

El coeficiente de dolarización del individuo sigue un proceso $AR(1)$ simple. En sí mismo, este exhibe algún grado de persistencia el cual depende del valor de a_i . Observe que un choque en ξ_{it} —una actualización de información o noticias— modifica x_{it} en impacto, en tanto esto cambia los pronósticos individuales i de R_i y en consecuencia, hace que los individuos revisen su cartera. Si $\xi_{it} > 0$ ($\xi_{it} < 0$), x_{it} aumenta (disminuye) puesto que el exceso de rendimientos esperados de mantener dólares es más alto (bajo). *Ceteris paribus*, en periodos futuros, la trayectoria del valor esperado de R_i se ajusta suavemente hacia su valor de largo plazo y asimismo, se ajusta la participación de los dólares en la cartera de inversión.

La dinámica de las decisiones de dolarización individuales refleja que con información ruidosa sobre rendimientos, los individuos tienen que confiar en información pasada para pronosticarlos en forma óptima y por tanto deben reaccionar con cautela ante las noticias. En la medida que sus decisiones de cartera pasadas contienen información histórica sobre los rendimientos, para los individuos resulta óptimo hacer que sus coeficientes de dolarización dependan de los coeficientes de dolarización históricos.²²

óptimo en el sentido que minimiza el EMC, un período hacia adelante, mientras que el ES es básicamente *ad hoc*. Este tema se discute con más detalle en la sección 5.

²² Un resultado similar pero en un modelo diferente se encuentra en Aoki (2003). En este documento, el banco central fija tasas de interés en un entorno de información ruidosa sobre la producción e inflación. La regla

3.4. Dinámica agregada

En un mundo estático, los efectos de la agregación son bien conocidos: tiende a alisar los movimientos erráticos individuales y a rellenar las discontinuidades que puedan estar presentes a nivel desagregado. Dentro de un marco dinámico, la agregación también aumenta la persistencia.²³ Para ver porque ocurre esto, considere un grupo de individuos que mantienen una pequeña cantidad de activos en dólares y enfrentan un choque agregado que los hace más atractivos (es decir, una fuerte depreciación real). Según (9), estos individuos aumentarán su tenencia de dólares inmediatamente. Pero después, este grupo tendrá que revisar sus expectativas sobre rendimientos futuros a favor del activo denominado en dólares; por ese motivo, se perpetúa el efecto del choque en la dolarización agregada. Así, debido a la falta de información perfecta, resumida en la ecuación (9), la persistencia moderada en la formación de la cartera individual se exagera debido a la agregación.²⁴

Se tiene que X_t es el coeficiente de dolarización de la economía. En el apéndice B se muestra que la agregación (9) a través de la distribución de a se convierte en el proceso siguiente.

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

donde los A_s ($s = 1, 2, \dots$) son coeficientes, \tilde{M} es una constante y \tilde{U}_t es una perturbación agregada serialmente no correlacionada. Como se sugirió antes, un resultado notable es que siendo la dolarización a nivel individual un proceso $AR(1)$, ésta se vuelva un proceso $AR(\infty)$ luego de la agregación –proceso usualmente conocido como de memoria *larga*.

Como Lewbel (1994) enfatiza, los coeficientes en (10) están estrechamente relacionados con la forma de $F(a)$. En el

de política óptima implica alguna persistencia proveniente de la prudencia que demanda la falta de información perfecta.

²³ La referencia clásica para la econometría de este efecto es Granger (1980), quien asume que $F(a)$ es una distribución beta. Ver también Pesaran (2003) y Zaffaroni (2004) para desarrollos recientes.

²⁴ Ver Michelacci (2004) para un análisis similar.

apéndice B se muestra como estos coeficientes satisfacen la siguiente recursión.

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

donde m_s es el s -ésimo momento de la distribución de a , $m_s = \int a^s dF(a)$. Por lo que resulta fácil verificar que:

$$\begin{aligned} \text{promedio } (a) &= m_1 = A_1 \\ \text{varianza } (a) &= m_2 - m_1^2 = A_2 \\ \text{sesgo } (a) &= (m_3 - 3m_1m_2 + 2m_1^3)(m_2 - m_1^2)^{-3/2} = (A_3 - A_1A_2)(A_2)^{-3/2} \end{aligned}$$

Estas relaciones nos permiten determinar como la distribución de las capacidades de predicción afectan la persistencia a nivel agregado. Cuanto mayor es A_1 , mayor es el promedio, lo que implica que el individuo *promedio* tiene él mismo un comportamiento más persistente, esto se refleja en un X_t más persistente. Por otro lado y de manera sorprendente, un A_2 mayor produce más persistencia: entre mayor sea la heterogeneidad entre los individuos, será más persistente el coeficiente de dolarización agregada. Finalmente, como señala Zaffanori (2004), el comportamiento de baja frecuencia del agregado está determinado por la forma de la distribución transversal cuando $a \rightarrow 1^-$. Entonces una distribución con una cola izquierda ancha ($A_3 < A_1A_2$) que indica una mayor cantidad de individuos persistentes ($a \approx 1$) podría sugerir una mayor persistencia agregada.

Es claro ahora que este marco puede probarse de forma directa. Si las estimaciones de A_s que usan datos agregados, son inconsistentes con la noción de diversos procesos dinámicos agregados en (10), entonces tendríamos que rechazar el modelo.²⁵ Los síntomas más obvios de contradicción serían una estimación no positiva de A_2 , la varianza de $F(a)$ ²⁶ o un valor muy negativo para A_1 , el promedio.

Es importante tener en cuenta que, la cantidad de información sobre el comportamiento individual que puede ser

²⁵ O las suposiciones detrás de la agregación. Ver apéndice B.

²⁶ Note que $A_2 = 0$, implica una distribución degenerada de a en el punto A_1 , es decir, un modelo con un agente representativo o individuos idénticos.

inferido a partir de los datos agregados es sin duda limitada. Distintos supuestos sobre las decisiones individuales pueden resultar consistentes con una variable agregada observada. Sin embargo, hay algunos hechos que se muestran en la siguiente sección, que sustentan las principales hipótesis de este documento y las predicciones del modelo teórico.

4. Evidencia empírica

En esta sección probamos si la dinámica del coeficiente de dolarización en los países seleccionados puede considerarse como resultado de la agregación de los depositantes heterogéneos. En otras palabras, se estiman los parámetros A_s , en la ecuación (10) y se investiga, a partir de los momentos estimados de la distribución subyacente $F(a)$, que tan heterogéneos son los participantes en el mercado de depósitos en dólares.

4.1. Aspectos econométricos y datos

Cabe mencionar tres puntos antes de presentar algunos resultados. En primer lugar y sin sorpresa alguna, cada coeficiente de dolarización X_t que consideramos, presenta raíz unitaria.²⁷ Para evitar los sesgos que la existencia de raíces unitarias genera en la estimación, se estima (10) en primeras diferencias.

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

El apéndice B muestra que (12) no es sólo la primera diferencia de (10), sino también el resultado de agregar (9) después de obtener primeras diferencias. Por lo tanto, los coeficientes en (12) son en realidad los mismos que en (10). La perturbación U_t^\dagger , presenta autocorrelación y heteroscedásticidad²⁸ por lo que se necesita inferencia robusta.

En segundo lugar, debido a las limitaciones de los datos no es posible estimar la ecuación (12) tal como está especificada. El tamaño de muestra es finito, por lo tanto resulta inevitable tener que truncar los rezagos del proceso $AR(\infty)$.

Por último, cuando resulta conveniente, se consideran dinámicas aún más ricas que las sugeridas por nuestros muy utilizados modelos teóricos, introduciendo un componente

²⁷ Los resultados de las pruebas de la raíz unitaria están disponibles mediante solicitud de los autores. Ver también apéndice C.

²⁸ Ver Pesaran (2003) para más detalles.

$MA(1)$ en (12). En la práctica, este hecho no tiene otra implicación para nuestro análisis que permitir mejores estimaciones de A_3 . Según Lewbel (1994), cuando esta presente un componente MA sólo un número finito de momentos de $F(a)$ puede ser recuperado debido a que el componente autorregresivo

CUADRO 1. MODELOS ARIMA DEL COEFICIENTE DE DOLARIZACIÓN DE DEPÓSITOS EN PAÍSES SELECCIONADOS

<i>Modelo ARIMA</i>	A_1	A_2	A_3	A_4	θ	$A_3 - A_1A_2$	\bar{R}^2
México (1985.T4 a 2005.T3, N=77)							
(4,1,0)	0.221 ^a (0.078)	0.199 ^a (0.078)	-0.192 ^a (0.072)	0.114 ^b (0.064)		-0.236 ^a (0.095)	0.221
(4,1,1)*	0.480 ^a (0.111)	0.195 ^a (0.094)	-0.216 ^a (0.063)	0.251 ^a (0.047)	-0.097 ^a (0.018)	-0.310 ^a (0.086)	0.261
Perú (1980.T1 a 2005.T3, N=94)							
(2,1,0)*	0.173 ^a (0.063)	0.142 ^a (0.158)				-0.024 ^b (0.013)	0.200
(2,1,1)	0.186 ^b (0.094)	0.139 ^a (0.065)			-0.058 (0.143)	-0.026 (0.016)	0.173
Polonia (1985.T4 a 2002.T4, N=69)							
(2,1,0)*	0.474 ^a (0.016)	0.113 ^a (0.052)				-0.053 ^a (0.024)	0.215
(2,1,1)	0.476 ^a (0.010)	0.111 ^a (0.049)			-0.007 (0.045)	-0.053 ^a (0.024)	0.275
Uruguay (1985.T1 a 2005.T3, N=83)							
(2,1,0)	0.218 ^a (0.091)	0.290 ^a (0.116)				-0.063 ^a (0.029)	0.153
(2,1,1)*	0.265 ^b (0.147)	0.215 ^a (0.055)			-0.093 ^a (0.034)	-0.057 ^b (0.033)	0.196

NOTAS: Estimaciones de máxima verosimilitud. Las cifras entre paréntesis son errores estándar robustos (consistentes). ^{a, b} Denota un nivel de significancia de 5%(10%). El error estándar del tercer momento central $A_3 - A_1A_2$ se calculó con el método delta. \bar{R}^2 es el R^2 ajustado. Las regresiones incluyen una constante, en caso de ser necesario, además incluyen algunas variables *dummy* para la remoción de valores atípicos. En todas las ecuaciones mostradas, las pruebas Breusch-Godfrey y Jarque-Bera indicaron que los residuos son no correlacionados y están normalmente distribuidos. Las especificaciones preferidas están marcadas con a*.

infinito en X_t (o en ΔX_t) no puede separarse del parámetro MA , es decir θ . Esta es una apreciación sustantiva más teórica que empírica; como se precisó anteriormente, nuestra intención no es recuperar cada momento de $F(a)$ sino sólo los primeros.

Se utiliza información de Perú y Uruguay (dos países fuertemente dolarizados), México y Polonia. Los datos son de frecuencia trimestral y van aproximadamente desde mediados de los años ochenta hasta mediados de la década actual. Como es habitual en la literatura de dolarización, X_t se mide como el coeficiente de depósitos de moneda extranjera del sector privado en el sistema bancario nacional con respecto a $M2$.²⁹ Esta información está disponible y nuestras fuentes son las páginas Web de varios bancos centrales y la base de datos de Estadísticas Financieras Internacionales del Fondo Monetario Internacional. La regresión con las series de tiempo más cortas (Polonia) tiene $N = 69$ observaciones, mientras que la mayor (Perú) tiene $N = 94$.

4.2. Resultados

Para cada país primero se estima un modelo ARIMA (2,1,0). Luego, realizamos una prueba de autocorrelación en los residuos e incluimos rezagos adicionales hasta que los residuos fuesen no correlacionados serialmente. En cada caso, no se necesitaron más de 2 rezagos, excepto en México, en el que se requirieron 4 rezagos. Con la finalidad de garantizar robustez, incluimos además un componente MA en la mejor especificación autoregresiva. El cuadro 1, muestra para cada país los mejores modelos autoregresivos, ARIMA (2,1,1) o ARIMA (4,1,0) y las ecuaciones ARIMA (2,1,1) o ARIMA (4,1,1) correspondientes. La columna etiquetada como θ contiene el coeficiente MA estimado. Para cada país marcamos con una * la mejor especificación, es decir, el modelo más parsimonioso que describe suficientemente bien a los datos.

²⁹ Una popular definición alternativa del coeficiente de dolarización discrimina entre residentes y no residentes, que incluye depósitos de residentes en el extranjero (Ize y Levy Yeyati, 2003). No se usó esta definición en nuestro análisis empírico, debido a que las correspondientes series de tiempo disponibles son pequeñas para el grupo de países analizados.

Un hecho interesante que se muestra en el cuadro 1, es la similitud entre las estimaciones de Perú y Uruguay, y entre aquellos de México y Polonia. Hay que recordar que Perú y Uruguay son economías fuertemente dolarizadas (por encima de 50%), mientras que México y Polonia, si bien han mostrado coeficientes de dolarización considerables a mediados de los años noventa, cuentan con coeficientes de dolarización menores al 30%. Sin embargo, existen características comunes entre los cuatro países que se discutirán más adelante.

Un resultado robusto que es común a todos los países y a las especificaciones dentro del mismo país es que los coeficientes A_1 y A_2 son significativamente positivos. Se debe recordar que A_1 es el promedio de $F(a)$ que en particular no resultó ser cercano a uno en ninguno de los casos, y que A_2 es su varianza. Además, en Perú y Uruguay, los coeficientes son de magnitudes comparables $A_2 \approx A_1$ lo que significa que la distribución subyacente de $F(a)$ es bastante dispersa en ambos casos. Los términos a son bastante heterogéneos, A partir del modelo que se prefirió para Perú se tiene que $A_1 \pm 2\sqrt{A_2} \approx [-0.60, 0.92]$, mientras que en el caso de Uruguay el intervalo es aún más ancho $A_1 \pm 2\sqrt{A_2} \approx [-0.66, 1.20]$. Para el caso de México, $A_1 \pm 2\sqrt{A_2} \approx [-0.40, 1.36]$, mientras que en el de Polonia $A_1 \pm 2\sqrt{A_2} \approx [-0.20, 1.15]$. Estos estimados implican un coeficiente de variación, $\sqrt{A_2} / A_1$ de 2.18 para Perú, 1.75 para Uruguay, 0.91 para México y 0.71 para Polonia. Estos resultados muestran que las economías fuertemente dolarizadas parecen tener una distribución de $F(a)$ mucho más dispersa, lo que es consistente con la idea de costos de participación decrecientes a medida que la dolarización se expande. Sin embargo, aún en el caso de México y Polonia, (donde $A_1 > A_2$) la estimación de la heterogeneidad subyacente resultó ser elevada.

A partir de los intervalos antes descritos se puede observar que la mayor cantidad de depositantes tiene un comportamiento estacionario $|a| < 1$. Un resultado más controversial es que algunos individuos parecen tener dinámicas explosivas, $a > 1$, Sin embargo, esto no significa necesariamente que los mismos individuos mantienen $a > 1$ en todos los periodos, sino que en cada período de tiempo una cantidad de individuos muestra este tipo de comportamiento. El mismo argumento

se aplica al hecho de que algunos depositantes parecen tener un $a < 0$, que implica un patrón de dolarización individual, errático.

Las estimaciones del tercer momento central implicado $A_3 - A_1A_2$ en cada país indican que $F(a)$ está sesgada hacia la izquierda. Cuando $A_1 > 0$, podría esperarse una distribución $F(a)$ sesgada hacia la izquierda, si hubiera la mezcla de un punto de masa por encima del promedio (individuos relativamente persistentes, aquellos que cambian su cartera en forma lenta) y algunos individuos con un a cercano a cero (correspondiente a aquellos que cambian su cartera más rápido). Así, el sesgo negativo es consistente con mercados de depósitos en dólares en donde un experto financiero comparte el mercado con un no experto.

5. Una reinterpretación

Para verificar la robustez de nuestros resultados, en esta sección modificamos ligeramente el modelo teórico. En la sección 3, las diferencias entre los agentes estuvieron centradas en que tan rápido cada agente ajusta su pronóstico cuando recibe nueva información (y a partir de allí, que tan rápido cambia su cartera). Se argumentaba que estas diferencias provenían de las diferencias entre su capacidad para realizar predicciones. A continuación, nos concentramos en la heterogeneidad entre las capacidades de las personas para extraer señales útiles. Postulamos que los individuos enfrentan coeficientes de señales a ruido q_i idiosincrásicas. Esto racionaliza de una manera simple el hecho de que los agentes que tienen un elevado q_i , los expertos financieros pueden extraer más información del indicador ruidoso S_t que aquellos que poseen un bajo q_i , los no expertos. A diferencia de los no expertos, los expertos financieros pueden distinguir si los cambios en S_t reflejan cambios en R_t o si sólo es ruido.

5.1. Reformulación del modelo

Escribimos el sistema (4) como *modelo de nivel local*.³⁰

$$(13) \quad \begin{aligned} r_{it+1} &= r_{it} + \omega_{it+1} & \omega_{it} &\sim iid(0, \sigma_\omega^2), \\ S_t &= r_{it} + \epsilon_{it} & \epsilon_{it} &\sim iid(0, \sigma_{\epsilon_i}^2) \end{aligned}$$

en donde ahora permitimos que varíe la varianza del ruido, $\sigma_{\epsilon_i}^2$. El coeficiente de señal a ruido es $q_i = \sigma_w^2 / \sigma_{\epsilon_i}^2$ y desempeña un papel clave para determinar como se ponderan las observaciones ruidosas para la extracción de señales (y predicción). A medida que q_i es mayor, se descuentan más observaciones del pasado cuando se realiza el pronóstico del futuro.

Para conveniencia de presentación se define $\tilde{v}_{it} = v_{it} \sigma_{\epsilon_i}^{-2}$. El esquema de actualización del ECM del \hat{r}_{it} —antes ecuación (7)—y su valor de equilibrio se transforma en:

³⁰ Ver Harvey (1989, cap. 4) y nota de pie 21.

$$(14) \quad \tilde{v}_{it+1} = \frac{\tilde{v}_{it}(1+q_i) + q_i}{\tilde{v}_{it} + 1} \quad \text{y} \quad \tilde{v}_i = \frac{q_i + \sqrt{q_i^2 + 4q_i}}{2},$$

mientras que la regla de pronóstico –antes (5)– ahora puede ser escrita como.

$$(15) \quad \hat{r}_{it+1} = (1-k_i)\hat{r}_{it} + k_i S_t, \quad \text{where } k_i = \frac{\tilde{v}_i}{\tilde{v}_i + 1}.$$

Dada la ecuación (15), la inversión en dólares óptima sigue el proceso:

$$(16) \quad x_{it} = (1-k_i)x_{it-1} + \frac{1}{\tilde{v}_{it} + 1} S_t.$$

Es fácil mostrar que k_i es creciente en q_i ,³¹ lo cual implica que aquellos individuos con elevado q_i pueden extraer más información de la señal en cada periodo. Más claramente, esto también implica que aquellos individuos tendrán coeficientes de dolarización menos persistentes. Como muestra la ecuación (16), a mayor k_i menor será el grado de persistencia de los coeficientes de dolarización. También cabe precisar que un cambio en señal S_t modifica en impacto al coeficiente de la dolarización, con un grado de reacción mayor de aquellos individuos que poseen más capacidad para extraer información.

5.2. Agregación y aspectos econométricos adicionales

Con un ligero abuso de notación, denotamos $a_i = 1 - k_i$. Recuerde $S_t = R_t + \epsilon_t$ donde ϵ_t es un choque agregado, Entonces, la agregación de (16) (ver apéndice B) lleva a:

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

Que a diferencia de (10) incluye un rezago distribuido de R_t . Esta diferencia nace como consecuencia de asumir distintas fuentes de heterogeneidad a nivel individual, y tiene una implicación directa sobre nuestro análisis: si la heterogeneidad

³¹ Ya que $\partial k_i / \partial q_i = \partial k_i / \partial \tilde{v}_i \cdot \partial \tilde{v}_i / \partial q_i$. con $\partial k_i / \partial \tilde{v}_i = (\tilde{v}_i + 1)^{-2} > 0$ y $\partial \tilde{v}_i / \partial q_i = 0.5 + (q_i + 2)(q_i^2 + 4q_i)^{-0.5} > 0$.

individual se aproxima mejor por medio de (13) que en lugar de (4), entonces las estimaciones en el cuadro 1 pueden estar sesgadas debido a la omisión de variables relevantes. A continuación, se extienden los modelos ARIMA del cuadro 1, para evaluar si esta presunta omisión cambia en alguna forma nuestras conclusiones principales.

Como se discutió en la sección 4.1, la ecuación final que se estimaría es:

$$(18) \quad \Delta X_t = \sum_{r=1}^{p_x} A_r \Delta X_{t-r} + \sum_{r=0}^{p_R} \beta_r \Delta R_{t-r} + U_t^*$$

donde p_x y p_R son rezagos finitos. La presencia de R_t y sus rezagos en (18) se explica porque los individuos en el modelo teórico basan sus decisiones exclusivamente en esta variable. No obstante, un marco de modelación más rico puede entender con facilidad la ecuación (18) a:

$$(19) \quad \Delta X_t = \sum_{r=1}^{p_x} A_r \Delta X_{t-r} + \sum_{r=0}^{p_R} \beta_r^D \Delta R_{t-r}^D + \sum_{r=0}^{p_R} \beta_r^P \Delta R_{t-r}^P + U_t^*$$

Como $R_t = R_t^D - R_t^P$, la ecuación (19) incluye a la ecuación (18) que es una versión restringida de $\beta_r^D = -\beta_r^P$ para cada r . Por eso, de ahora en adelante utilizaremos (19).

Un tema empírico que surge con la introducción de rendimientos reales en las ecuaciones agregadas, es precisamente, como medirlos. Los rendimientos *verdaderos* involucran las expectativas de variables macroeconómicas futuras, de las cuales no existe información disponible para los países analizados. Se denotan las tasas de interés nominal i_t^P e i_t^D en moneda nacional y en USD respectivamente, la depreciación nominal δ_t (es decir, el cambio porcentual en el tipo de cambio nominal, la moneda nacional por USD) y π_t , la inflación del IPC. Se consideran dos medidas alternativas para los rendimientos reales.

<i>ex ante:</i>	<i>ex post:</i>
$R_t^P = \frac{1 + i_t^P}{1 + \pi_{t-1}} - 1$	$R_t^P = \frac{1 + i_t^P}{1 + \pi_t} - 1$
$R_t^D = \frac{(1 + i_t^D)(1 + \delta_{t+1})}{1 + \pi_{t+1}} - 1$	$R_t^D = \frac{(1 + i_t^D)(1 + \delta_{t+1})}{1 + \pi_t} - 1$

Los datos del IPC y del tipo de cambio nominal están disponibles. Para i_t^P , se utiliza la tasa pasiva en moneda nacional para Perú, Polonia y Uruguay, y la tasa de ahorro en moneda nacional para México. Para i_t^D , se usan datos de tasas de interés pagadas sobre depósitos nacionales en dólares sólo en el caso de Perú y Uruguay. Para México y Polonia el i_t^D se aproxima a

CUADRO 2. ECUACIONES AUMENTADAS

	A_1	A_2	A_3	$A_3 - A_1A_2$	$H_0: \beta = 0$	p_R	\bar{R}^2
México (1985.T4 a 2005.T3, N=77)							
<i>ex ante</i>	0.291 ^a (0.096)	0.202 ^a (0.066)	-0.273 ^a (0.092)	-0.331 ^a (0.113)	11.50 ^a [0.000]	2	0.554
<i>ex post</i>	0.129 (0.108)	0.287 ^a (0.089)	-0.240 ^a (0.080)	-0.278 ^a (0.091)	22.56 ^a [0.000]	2	0.565
Perú (1980.T1 a 2005.T3, N=89)							
<i>ex ante</i>	0.242 ^a (0.043)	0.195 ^a (0.047)	0.003 (0.053)	-0.047 ^a (0.015)	9.086 ^a [0.000]	3	0.435
<i>ex post</i>	0.501 ^a (0.098)	0.138 ^b (0.083)	-0.027 (0.068)	-0.069 ^b (0.036)	30.85 ^a [0.000]	2	0.649
Polonia (1985.T4 a 2002.T4, N=68)							
<i>ex ante</i>	0.449 ^a (0.043)	0.132 ^a (0.058)	-0.002 (0.049)	-0.059 ^a (0.022)	1.638 [0.203]	3	0.275
<i>ex post</i>	0.586 ^a (0.077)	0.164 ^a (0.070)	-0.123 (0.160)	-0.096 ^a (0.043)	2.402 ^b [0.099]	4	0.394
Uruguay (1985.T1 a 2005.T3, N=80)							
<i>ex ante</i>	0.252 ^a (0.104)	0.280 ^a (0.114)	-0.109 (0.140)	-0.070 ^b (0.038)	3.153 ^a [0.049]	3	0.124
<i>ex post</i>	0.267 ^a (0.103)	0.349 ^a (0.117)	-0.073 (0.143)	-0.093 ^b (0.047)	2.189 [0.119]	2	0.152

NOTAS: Estimaciones de mínimos cuadrados bietápicos. Los instrumentos para el R_t^D y R_t^P (y para el *ex ante* R_{t-1}^D y R_{t-1}^P) son cambios de los precios del petróleo. El crecimiento del PIB en los Estados Unidos y los valores rezagados de estas y las variables de R. Las cifras entre paréntesis son errores estándar robustos (consistentes).^{a, b} Denota un nivel de significancia de 5% (10%). Las cifras en la columna $H_0: \beta = 0$ son estadísticos F y los valores p se muestran entre corchetes. El tamaño del rezago p_R fue escogido para minimizar el criterio de Schwarz. Para Perú, Polonia y Uruguay, se fija $A_3 = 0$ para computar el tercer momento central y su desviación estándar. Las pruebas de diagnóstico sugirieron residuos bien comportados. Todas las regresiones incluyen una constante.

la tasa de depósito de los Estados Unidos.³² Nuestras fuentes son los bancos centrales y las estadísticas financieras internacionales del FMI.

Finalmente la presencia de un rendimiento contemporáneo (19) puede hacer más posible el sesgo de endogeneidad. Usamos el procedimiento de mínimos cuadrados bivariados para estimar esta ecuación. Los instrumentos empleados se detallan más adelante en el cuadro 2. Cabe mencionar que los mínimos cuadrados ordinarios o la exclusión de los rendimientos contemporáneos no alteraron los resultados principales de esta revisión de robustez.³³

5.3. Resultados

El cuadro 2 presenta los resultados de la estimación. Por simplicidad, se omitieron los coeficientes de los rendimientos (ya que estos no son de interés directo para nuestro análisis) pero sí se informa el estadístico- F que evalúa la significancia total de estos coeficientes. Establecimos un rezago de tamaño $p_x = 3$. Esta es la mejor elección para México, para los demás países, el óptimo es $p_x = 2$, pero se mantiene $p_x = 3$ para garantizar que se consideren todos los efectos autoregresivos. La elección de p_x que se incluye en el cuadro, responde a la minimización del criterio de Schwarz.

Recuerde que, mediante la estimación de las ecuaciones aumentadas se está evaluando si los resultados del cuadro 1 son robustos. Entonces, ¿lo son? En general, así es. Una comparación rápida entre las estimaciones del cuadro 2 con las del cuadro 1 revela que debido a la presencia de los rendimientos, el ajuste de varias ecuaciones aumenta, pero las estimaciones

³² Desafortunadamente, no se encontraron series de tiempo lo bastante largas de la prima por riesgo país para tener una mejor medida del R_t^D en estos dos países. Sin embargo, los resultados de la estimación no cambiaron cuando se consideró la tasa LIBOR (en USD, a varios plazos) en lugar de la tasa pasiva de los Estados Unidos.

³³ No se encuentra ninguna relación de cointegración significativa entre: X_t , R_t^P y R_t^D o entre X_t y R_t para tratar (18) o (19) como un modelo de corrección de errores. Existen algunos quiebres estructurales en nuestro intervalo de 20 años de datos que pueden explicar esta falla. De manera consistente a esto, los niveles de los rendimientos no parecen tener suficiente poder explicativo en las ecuaciones del cuadro 2.

de A_1A_2 y $A_3 - A_1A_2$ no cambian mucho. La notable excepción para este patrón es el caso mexicano, cuando los rendimientos son medidos de manera *ex post*, debido a que A_1 pierde significancia estadística. Sin embargo, la conclusión principal de las secciones anteriores aún se mantiene, cualitativa y también cuantitativamente: la heterogeneidad, de los tomadores de decisiones que subyace a los coeficientes de dolarización agregados es alta y este hecho conlleva a la persistencia de la dolarización agregada.

6. Comentarios finales

En países con coeficientes de dolarización elevados, la participación del mercado de depósitos en dólares se ha vuelto masiva. La desregulación, innovación y liberación financiera, así como el crecimiento de los mercados de cambios informales han facilitado la participación de grupos de agentes bastante heterogéneos en el mismo mercado –desde una gran empresa que utiliza modernas técnicas de manejo de cartera hasta personas poco informadas quienes basan sus decisiones de cartera simplemente en su experiencia y en información limitada–. Este documento muestra que dicha heterogeneidad, resulta ser suficiente para generar la persistencia de los coeficientes de dolarización agregada. La evidencia empírica de los tres países de América Latina y de Polonia apoya esta hipótesis.

La heterogeneidad en las decisiones de dolarización individual tiene interesantes implicaciones de política. Ize y Levy Yeyati (2003) concluyen razonablemente que una condición necesaria y suficiente para la desdolarización es una mayor flexibilidad del tipo de cambio. En nuestro modelo, esta condición no es suficiente (aunque si consideramos que es necesaria) ya que puede existir una gran cantidad de individuos que no respondería del todo a esa mayor volatilidad. Este resultado implica que es necesaria una política más activa de comunicación por parte del banco central, que transmita mejor su política de tipos de cambio más flexibles, y su compromiso con la estabilidad de precios a un conjunto de agentes más amplio, especialmente a los menos informados. De esta manera, los formuladores de la política monetaria estarían contribuyendo a reducir la heterogeneidad individual y así la persistencia agregada.

Esta implicación de política es en particular relevante para las economías en desarrollo con un régimen de metas explícitas de inflación y para los que están evaluando adoptar este tipo de régimen, debido a que se apoya fuertemente en estrategias de transparencia y comunicación. Nuestros análisis sugiere que los beneficios de dicho régimen de política en

cuanto a la reducción de la dolarización podrían llegar a ser limitados, a menos que el banco central comunique en forma efectiva las implicaciones y beneficios de dicho régimen a los participantes menos informados en el mercado de cambios.

Apéndices

A. La distribución de dotaciones y capacidades

Nuestros resultados asumen que los agentes son homogéneos en sus dotaciones de riqueza. En particular, el análisis se restringe al caso en el que cada agente tiene una dotación de tamaño uno. En este apéndice se muestra que los resultados no dependen de este supuesto y los resultados se mantienen para un caso más general, uno en que los agentes tienen niveles de riqueza diferentes, pero donde la distribución de sus capacidades (a) entre los agentes está positivamente correlacionada con su riqueza. Un supuesto factible de observarse en la realidad.

Considerar la ecuación (9). Por simplicidad, se fija $\mu = 0$ y se asume que el ingreso agregado es igual a uno y que existen dos agentes en la economía: uno con capacidad a_1 e ingreso n_1 y la otra con capacidad a_2 e ingreso $n_2 = 1 - n_1$. Entonces:

$$(A1) \quad (1 - a_i L)x_{it} = a_i \left(\frac{1}{v_i + \sigma_\epsilon^2} \right) \xi_{it}, \quad \text{para todo } i = 1, 2.$$

Después de generar un polinomio de rezago común para ambos procesos, se tiene:

$$(A2) \quad \begin{aligned} (1 - a_2 L)(1 - a_1 L)x_{1t} &= a_1 \left(\frac{1 - a_2 L}{v_1 + \sigma_\epsilon^2} \right) \xi_{1t} \\ (1 - a_1 L)(1 - a_2 L)x_{2t} &= a_2 \left(\frac{1 - a_1 L}{v_2 + \sigma_\epsilon^2} \right) \xi_{2t}. \end{aligned}$$

El nivel de depósito en dólares, que coincide con el coeficiente de dolarización agregado, es $X_t = n_1 x_{1t} + n_2 x_{2t}$. Agregando las ecuaciones en (A2), se obtiene:

$$(A3) \quad (1 - a_1 L)(1 - a_2 L)X_t = n_1 a_1 \left(\frac{1 - a_2 L}{v_1 + \sigma_\epsilon^2} \right) \xi_{1t} + n_2 a_2 \left(\frac{1 - a_1 L}{v_2 + \sigma_\epsilon^2} \right) \xi_{2t}.$$

Define $\tilde{\xi}_{it} = n_i a_i (v_i + \sigma_\epsilon^2)^{-1} \xi_{it}$ para $i = 1, 2$. Entonces la ecuación (A3) al final de todo es:

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

Se tiene que el coeficiente de dolarización agregada sigue un proceso ARMA (2,1). Este ejemplo simple puede ser generalizado al caso de N procesos AR (1) (es decir, N distintos niveles de dotación y de capacidad); en tal caso el coeficiente de dolarización agregado sigue un proceso ARMA (N^* , N^*-1) donde $N^* \leq N$. El número de agentes involucrados se puede incrementar simplemente replicando un número arbitrario de veces el comportamiento individual de un agente con una capacidad a dada. Por lo tanto, los resultados de agregación derivados en el apéndice B, pasan por el supuesto de que la distribución de dotaciones está correlacionada con aquella de las capacidades de procesar información. Cuando $N \rightarrow \infty$, se obtiene el caso límite presentado en el apéndice C. Las derivaciones de este apéndice aplican la versión alternativa de nuestro modelo que lleva a la ecuación (16).

B. Agregación

Las presentes derivaciones siguen de cerca a Lewbel (2004). Para facilitar la notación desechamos el superíndice i , en este apéndice.

B.1. Ecuaciones (10) y (11)

Considere la ecuación (9),

$$(B1) \quad x_t = ax_{t-1} + m + u_t,$$

donde $m = \mu(1-\alpha)v^{-1}$ y $\mu_t = a(v + \sigma_\epsilon^2)^{-1} \xi_t$ dependen de a . Puesto que por construcción ξ_t es una secuencia de choques no correlacionados serialmente es decir, u_t . Sin embargo, u_t está correlacionado entre individuos. Sea \mathcal{E} el operador de expectativas a través de los individuos, $\mathcal{E}[z] = \int z dF(a)$, tal que $X_t = \mathcal{E}[x_t]$, $M = \mathcal{E}[m]$ y $U_t = \mathcal{E}[u_t]$. Luego de agregar (B1) se obtiene:

$$(B2) \quad X_t = \mathcal{E}[ax_{t-1}] + M + U$$

Definimos una variable aleatoria α_s , un escalar $A_s = \mathcal{E}[\alpha_s]$,

una recursión $\alpha_{s+1} = (\alpha_s - A_s)a$ con la siguiente condición inicial $\alpha_1 = a$. Note que para $s > 1$ la recursión previa implica que $\alpha_s = a^s - \sum_{r=1}^{s-1} a^{s-r} A_r$. Luego de aplicar el operador esperanza \mathcal{E} , obtenemos la ecuación (11) en la ecuación, donde $m_s = \mathcal{E}[a^s]$ es el *ese-ésimo* momento de la distribución de a , note también que:

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

donde $\text{cov}(\alpha_s, m)$ es la covarianza transversal de α_s y m , que no varía en el tiempo. Por otro lado, $\text{cov}(\alpha_s, u_{t-s})$ es la covarianza transversal de α_s y u_{t-s} que es dependiente del tiempo, pero como esta dependencia proviene de ξ_t , ésta no es serialmente correlacionada.

La ecuación (B3) muestra la recursión entre $\mathcal{E}[\alpha_s x_{t-s}]$ y $\mathcal{E}[\alpha_{s+1} x_{t-(s+1)}]$, luego de resolver.

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

Sea $V_t = \sum_{r=1}^{\infty} \text{cov}(\alpha_r, u_{t-r})$ y $\tilde{V} = E[V_t]$, donde E es el operador de expectativa, a través del tiempo. Definimos también $\tilde{M} = M + \sum_{r=1}^{\infty} \text{cov}(\alpha_r, m) + \tilde{V}$ y $\tilde{U}_t = U_t + V_t - \tilde{V}$. Después de reemplazar (B4) en (B2), obtenemos la ecuación (10) del texto principal $X_t = \sum_{r=1}^{\infty} A_r X_{t-r} + \tilde{M} + \tilde{U}_t$ donde \tilde{U}_t es serialmente no correlacionado.³⁴ Los supuestos detrás de la ecuación agregada son tales que, \tilde{M} y V_t son ambos finitos o las secuencias $\{\text{cov}(\alpha_s, m)\}_{s=1}^{\infty}$ y $\{\text{cov}(\alpha_s, u_{t-s})\}_{s=1}^{\infty}$ son absolutamente sumables.

B.2. Ecuación (12)

Considere ahora la ecuación (B1) en primeras diferencias:

$$\text{(B5)} \quad \Delta x_t = a\Delta x_{t-1} + u_t - u_{t-1},$$

³⁴ Aunque Pesaran (2004) muestra que es heteroscedástico.

de modo que después de la agregación $\Delta X_t = \mathcal{E}(a\Delta X_{t-r}) + U_t - U_{t-1}$. Siguiendo el mismo procedimiento aplicado para obtener la ecuación (B4).

$$(B6) \quad \mathcal{E}(a\Delta x_{t-1}) = \sum_{r=1}^{\infty} A_r \Delta X_{t-r} + V_t - V_{t-1},$$

de modo que ΔX_t pueda escribirse como:

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

que corresponde a la versión de la primera diferencia de la ecuación (10). El nuevo error agregado U_t^\dagger , no presenta correlación serial y los coeficientes son los mismos que aquellos en (10).

B.3. Ecuación (17)

Todos los resultados derivados anteriormente en este apéndice se cumplen para sus correspondientes análogos agregados. Note que esta ecuación puede escribirse como $x_t = ax_{t-1} + bS_t$, donde a y b son coeficientes individuales específicos mientras que S_t es un número agregado. Así, $\mathcal{E}[bS_t] = \mathcal{E}[b]S_t$, por lo tanto, $X_t = \mathcal{E}[ax_{t-1}] + \mathcal{E}[b]S_t$. La ecuación (B4) es ahora:

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

Denotamos $\beta_0 = \mathcal{E}[b]$, $\beta_r = \text{cov}(\alpha_r, b)$, $\hat{U}_t = \sum_{r=0}^{\infty} \beta_r \varepsilon_{t-r}$ y recordamos que $S_t = R_t + \varepsilon_t$. Mediante una simplificación mecánica de esta expresión obtenemos (17). El residuo agregado \hat{U}_t está serialmente correlacionado.

C. Una breve nota sobre integración financiera

Considere el siguiente modelo dinámico univariado:

$$(C1) \quad \Phi(L)(1-L)^d X_t = \Theta(L)\eta_t,$$

donde L es el operador de rezago, $\eta_t \sim iid(0, \sigma_\eta^2)$ y d es el

parámetro de diferenciación. Cuando $d=0$, X_t es estacionario y sigue un proceso ARMA $\Phi(L)X_t = \theta(L)\eta_t$. Cuando $d=1$, X_t tiene una raíz unitaria y de allí sigue un proceso ARIMA $\Phi(L)\Delta X_t = \theta(L)\eta_t$. Cuando d toma valores no enteros, se dice que X_t es un proceso ARMA (ARFIMA) fraccionalmente integrado. Cuando $d \in (0, 0.5]$, la función autovarianza de X_t desciende hiperbólicamente a cero haciendo de X_t un proceso estacionario de larga memoria. Para $d > 0.5$, X_t es no estacionario (tiene varianza infinita).

Granger (1980) ha mostrado bajo ciertos supuestos sobre $F(a)$ la –distribución de coeficientes autoregresivos individuales– la agregación del proceso AR (1) como (9) conduce a (C1).³⁵ En nuestra aplicación empírica, simplemente se fija $d=1$. Si el verdadero modelo se caracteriza por $d < 1$, los datos estarían sobre diferenciados, con posibles efectos negativos, en nuestra inferencia estadística.

CUADRO B1. PARÁMETRO DE INTEGRACIÓN FRACCIONARIA ESTIMADO EN COEFICIENTES DE DOLARIZACIÓN

	\hat{d}	$H_0: d = 0$		$H_0: d = 1$	
		Estadístico t	Valor p	Estadístico t	Valor p
México	0.825	2.376	0.0491	0.505	0.6294
Perú	0.932	3.883	0.0037	0.282	0.7843
Polonia	0.955	4.605	0.0025	0.219	0.8333
Uruguay	0.788	2.485	0.0378	0.667	0.5236

NOTAS: Se usa el método de estimación de Geweke y Porter-Hudak (conocido como GPH). El error estándar asintótico de \hat{d} es $\pi^2/6$, que es usado para computar los estadísticos t y los valores p . Ambas pruebas ($H_0: d = 0$ y $H_0: d = 1$) son de dos colas. Ver Baillie (1996) para una revisión de la modelación ARFIMA y para las críticas del estimador GPH.

El cuadro B1 muestra las estimaciones de d y pruebas: $H_0: d = 0$ y $H_0: d = 1$. No encontramos suficiente evidencia para rechazar $H_0: d = 1$, mientras que, $H_0: d = 0$ se rechaza sistemáticamente.

³⁵ Ver también Baillie (1996) y Zafarroni (2004).

Referencias

- Agénor, Pierre-Richard, y Nadeem U. Haque (1996), “Macroeconomic Management with Informal Financial Markets”, *International Journal of Finance and Economics*, vol. 1, n° 2, pp. 87-101.
- Aoki, Kosuke (2003), “On the Optimal Monetary Policy Response to Noisy Indicators”, *Journal of Monetary Economics*, vol. 50, n° 3, pp. 501-23.
- Baillie, Richard (1996), “Long Memory Processes and Fractional Integration in Econometrics”, *Journal of Econometrics*, vol. 73, n° 1, pp. 5-59.
- Calvo, Guillermo A., y Carmen M. Reinhart (2002), “Fear of Floating”, *Quarterly Journal of Economics*, vol. 107, n° 2, pp. 379-408.
- Céspedes, Luis F., Roberto Chang y Andrés Velasco (2004), “Balance Sheets and Exchange Rate Policy”, *American Economic Review*, vol. 94, n° 4, pp. 1183-1193.
- Cook, David (2004), “Monetary Policy in Emerging Markets: Can Liability Dollarization Explain Contractionary Devaluations?”, *Journal of Monetary Economics*, vol. 51, n° 6, pp. 1155-1181.
- De Nicoló, Gianni, Patrick Honohan y Alain Ize (2005), “Dollarization of Bank Deposits: Causes and Consequences”, *Journal of Banking and Finance*, vol. 29, n° 7, pp. 1697-1727.
- Evans, George W., y 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*, vol. 14, n° 2, pp. 227-238.
- Guidotti, Pablo E., y Carlos A. Rodríguez (1992), “Dollarization in Latin America: Gresham’s Law in Reverse?”, *IMF Staff Papers*, vol. 39, n° 3, pp. 518-544.
- Harvey, Andrew C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.
- Harvey, Andrew C., y Giuliano De Rossi (2006), “Signal extraction”, cap. 31, en Terence C. Mills and Kerry Patterson (eds.), *Palgrave Handbook of Econometrics*, Palgrave Macmillan.

- Ize, Alain, y Eduardo Levy Yeyati (2003), “Financial Dollarization”, *Journal of International Economics*, vol. 59, n^o 2, pp. 323-47.
- Kamin, Steven B., y Neil R. Ericsson (2003), “Dollarization in Post-Hyperinflationary Argentina”, *Journal of International Money and Finance*, vol. 22, n^o 2, pp. 185-211.
- Levy Yeyati, Eduardo (2006), “Financial Dollarization: Evaluating the Consequences”, *Economic Policy*, forthcoming.
- Lewbel, Arthur (1994), “Aggregation and Simple Dynamics”, *American Economic Review*, vol. 84, n^o 4, pp. 905-918.
- Ljungqvist, Lars y 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*, vol. 51, n^o 7, pp. 1321-1352.
- Morón, Eduardo, y Diego Winkelried (2005), “Monetary Policy Rules for Financially Vulnerable Economies”, *Journal of Development Economics*, vol. 76, n^o 1, pp. 23-51.
- Pesaran, M. Hashem (2003), “Aggregation of Linear Dynamic Models: An Application to Life-Cycle Consumption Models under Habit Formation”, *Economic Modelling*, vol. 20, n^o 2, pp. 383-415.
- Quispe, Zenón (2000), “Monetary Policy in a Dollarized Economy: the Case of Peru”, en Mahadeva, Lavan and Gabriel Sterne (eds.), *Monetary Policy Frameworks in a Global Context*, Routledge & Bank of England, Londres-Nueva York.
- Reinhart, Carmen M., Kenneth S. Rogoff y Miguel A. Savastano (2003), *Addicted to Dollars*, NBER (Working Paper, n^o 10015).
- Savastano, Miguel A. (1996), “Dollarization in Latin America. Recent Evidence and some Policy Issues”, en 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*, vol. 21, n^{os} 2/3, pp. 391-416.
- Uribe, Martín (1997), “Hysteresis in a Simple Model of Currency Substitution”, *Journal of Monetary Economics*, vol. 40, n^o 1, pp. 185-202.
- Zaffaroni, Paolo (2004), “Contemporaneous Aggregation of Linear Dynamic Models in Large Economies”, *Journal of Econometrics*, vol. 120, n^o 1, pp. 75-102.

WHY DOLLARIZATION
IS SO PERSISTENT?

Paul Castillo and Diego Winkelried

Why dollarization is so persistent?

CENTRAL BANK AWARD “RODRIGO GÓMEZ, 2006”

1. Motivation

Even though dollarization is a relatively new research area, the experiences of many Latin American and transition economies during the 1990s have 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, USD) 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* behaviour 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 the 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

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 Ericson (2003).

consistent with the strong macroeconomic fundamentals observed in several highly dollarized countries (notably Peru and some transition economies in the early 2000s).

An alternative avenue to address this puzzle is to modify existing currency substitution models based on adjustment costs or network externalities. Guidotti and Rodriguez (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 in the dollarization ratio, they do not seem useful in advising policymakers how to dedollarize, as this may imply an implausible reduction in the knowledge stock.

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

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) extend 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

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 heterogeneous, 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 mark-ups than those in the banking sector.⁷ A typical *cambista* would hold a limited amount of money for business (say, between USD 3,000 and USD 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 take informed saving decisions.⁹

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

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

(a lot of) individuals can be taken roughly as having the same importance.

¹⁰The 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.

¹¹Our approach is related to other branches of the literature. For instance,

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 policy recommendations. Derivations and complementary results are shown in the appendix.

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 dedollarization (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 2000s

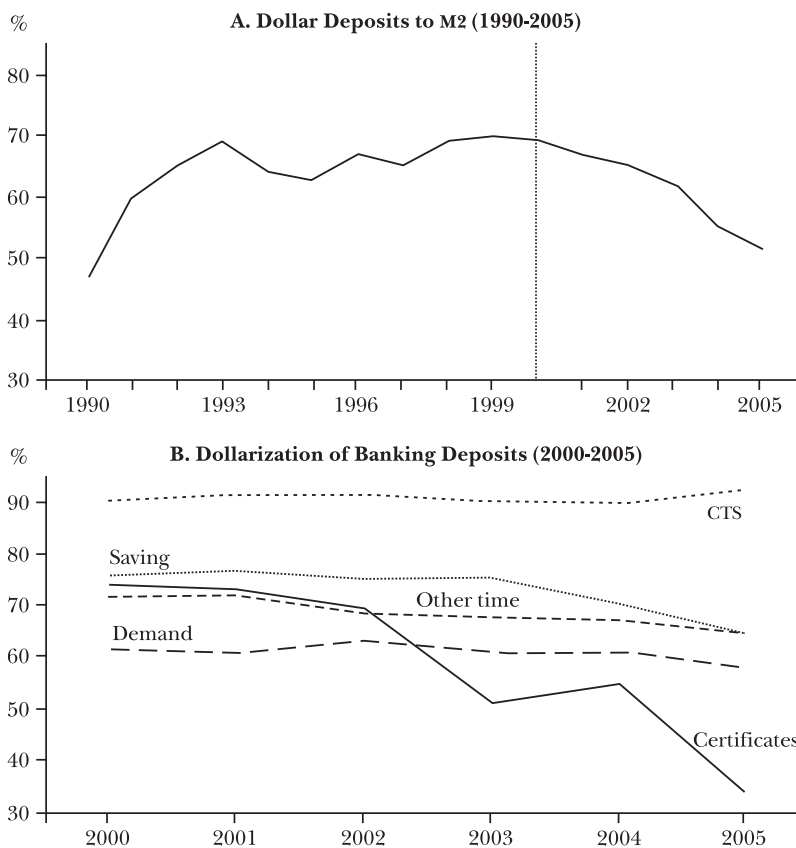
Although the Peruvian dollarization experience shares various of the aforementioned features, it has its own appeal.¹³ As shown in figure 1(A), in 1991 (after a four-digit hyperinflation in 1990) the ratio was 60% and has remained fluctuating roughly between 65% and 70% for a decade. Since 2000, it has 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.

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.

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 favourable 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

FIGURE 1. DEPOSIT DOLLARIZATION IN PERU, 1990-2005



SOURCE: Central Bank of Peru.

deposits dollars vis-à-vis holding deposits in domestic currency has fallen considerably in the early 2000s.

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 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.

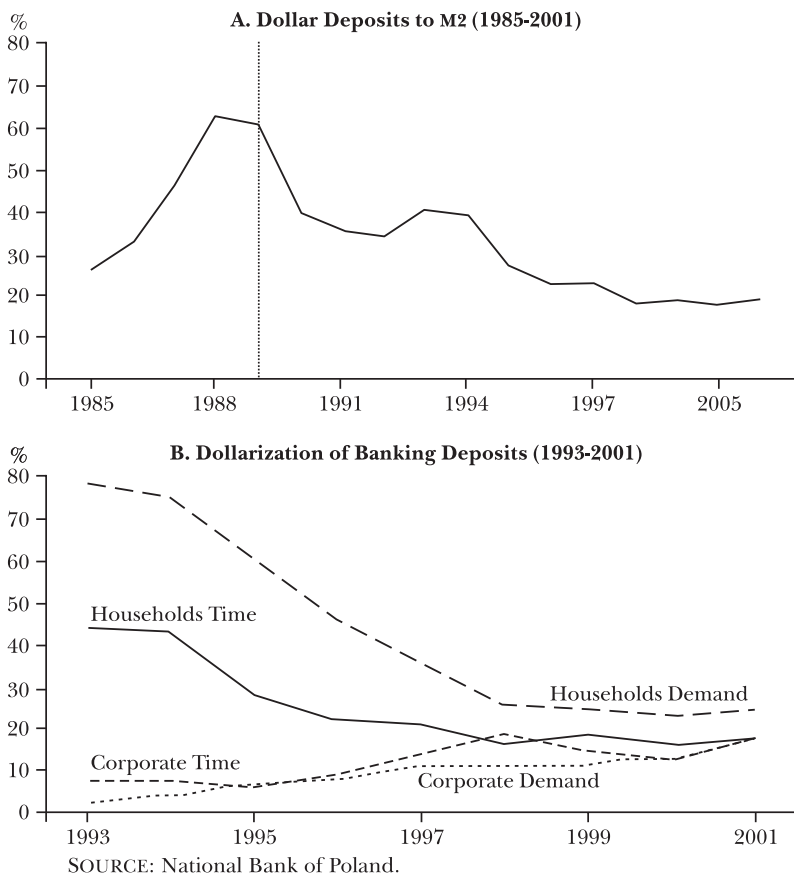
The figure also shows a moderate downward trend in the savings and other time deposits. About 80% of the saving and roughly half of the other time deposits are held by individuals. From 2001 to 2005 both ratios have decreased in about 10%. What is remarkable from figure 1(B) is the strong reaction of 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 has similar term than the CTS and the other time deposits, they are mainly held by firms and not by 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 1980s, 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

FIGURE 2. DEPOSIT DOLLARIZATION IN POLAND, 1985-2001



of a stabilization program (the so-called *shock-therapy*),¹⁴ dollarization ratios dropped to averages of 40% 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

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

stability. Dollarization decreased even more reaching by 2001 the level of 18%, comparable with that of developed European economies, as the United Kingdom.

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 four 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 neighbours 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 analyse how this translates into persistence.

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 share the same endowment, which is normalised 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 (*pesos* 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 pesos asset,

$$(1) \quad R_t = R_t^D - R^P.$$

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

$$(2) U_{it} = E_t [x_{it} R_{t+1}^D + (1 - x_{it}) R^P] - \left(\frac{1}{2}\right) \text{var}_t [x_{it} R_{t+1}^D + (1 - x_{it}) R^P]$$

¹⁵ Our analysis holds for agents with heterogeneous 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.

$$= E_t[x_{it}R_{t+1} + R^P] - \left(\frac{1}{2}\right) \text{var}_t[x_{it}R_{t+1}] = x_{it}\hat{r}_{it+1} + R^P - \left(\frac{1}{2}\right)(x_{it})^2 v_{it+1},$$

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 maximises (2) is:

$$(3) \quad x_{it} = \frac{\hat{r}_{it+1}}{v_{it+1}}.$$

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

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:

$$(4) \quad \begin{array}{ll} r_{it+1} & = \mu(1-a_i) + a_i r_{it} + \omega_{it+1} & \omega_{it} \sim iid(0, \sigma_\omega^2) \\ S_t & = r_{it} + \varepsilon_{it} & \varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2) \end{array}.$$

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_ω^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 $v_{it} = E[(r_{it} - \hat{r}_{it})^2]$ 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

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

where the forecasted value of r_{it} for next period is the projection of today's forecasted value plus a correction, an updating $\xi_{it} = \varepsilon_{it} + (r_{it} - \hat{r}_{it})$, 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 of the noisy indicator,

$$(6) \quad k_{it} = a_i \left(\frac{v_{it}}{v_{it} + \sigma_\varepsilon^2} \right).$$

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

$$(7) \quad v_{it+1} = \frac{v_{it}(a_i^2\sigma_\varepsilon^2 + \sigma_\omega^2) + \sigma_\varepsilon^2\sigma_\omega^2}{v_{it} + \sigma_\varepsilon^2}.$$

It is clear from equation that $v_{it+1} = f(v_{it})$. There is a fixed point such that $v_i = f(v_i)$ ¹⁹ and moreover, since $f'(v_i) < 1$ it is globally stable: regardless of the initial condition v_{i0} we have that $v_{it} \rightarrow v_i$ and consequently $k_{it} \rightarrow k_i = a_i v_i (v_i + \sigma_\varepsilon^2)^{-1}$ as $\tau \rightarrow \infty$. This means that as τ becomes larger, i.e. as each individual has performed the signal extraction exercise a

¹⁷ For instance, if the true DGP of R_t is a random walk, then the predictions of those with $a_i = 1$ 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.

¹⁹ The fixed point is the positive root of $v_i^2 + [(1-a_i^2)\sigma_\varepsilon^2 - \sigma_\omega^2]v_i - \sigma_\varepsilon^2\sigma_\omega^2 = 0$.

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 initialised 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:

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

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 $v_{it} \rightarrow v_i$, $k_{it} \rightarrow k_i$ and the optimal updating/forecasting rule (5), the optimal dollar investment (3) boils down to:

$$(8) \quad x_{it} = \frac{\hat{r}_{it+1}}{v_{it+1}} = \frac{\hat{r}_{it+1}}{v_i} = \frac{\mu(1-a_i) + a_i \hat{r}_{it} + k_i \xi_{it}}{v_i} = a_i x_{it-1} + \frac{\mu(1-a_i)}{v_i} + \left(\frac{k_i}{v_i} \right) \xi_{it}.$$

After plugging into, we get:

$$(9) \quad x_{it} = a_i x_{it-1} + \frac{\mu(1-a_i)}{v_i} + a_i \left(\frac{1}{v_i + \sigma_\epsilon^2} \right) \xi_{it}.$$

The individual's dollarization ratio follows a simple AR(1) process. As such, it exhibits some degree of persistence

²⁰ Convergence is monotonic ($v_{it} \geq v_{it+1} \geq v_i$) because v_{it+1} is based on more information than v_{it} . 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 method converges to the popular Exponential Smoothing method (ES) as $a \rightarrow 1$. However, the scheme explained here is optimal in the sense that it minimises the one step ahead MSE, whereas ES is basically ad hoc. We explore this issue further in section 5.

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 favour of the dollar asset, thereby perpetuating the impact effect of the shock on aggregate dollarization. Thus, the moderate persistence in the individual portfolio formation

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.

due to the lack of perfect information, summarised 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:

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

where the A_s ($s=1,2,\dots$) 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:

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

where m_s is the s -th *moment* of the distribution of a , $m_s = \int a^s dF(a)$. 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 behaviour, 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 behaviour of the aggregate is determined by the shape of the cross sectional distribution as $a \rightarrow 1^-$. Hence, a distribution with a

²⁴ See Michelacci (2004) for a similar analysis.

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 behaviour 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 root²⁷ and to avoid well-known biases in the estimation of autoregressive coefficients when a unit root is present we estimate (10) in first differences,

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

Appendix B shows that (12) is not only the first-differenced version of (10), but is also the result of aggregating 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 stylised 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

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

²⁸ See Pesaran (2003) for further details.

$F(a)$ 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

TABLE 1. ARIMA MODELS OF THE DEPOSIT DOLLARIZATION RATIO IN SELECTED COUNTRIES

<i>ARIMA model</i>	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 ^a (0.078)	0.199 ^a (0.078)	-0.192 ^a (0.072)	0.114 ^b (0.064)		-0.236 ^a (0.095)	0.221
(4,1,1)*	0.480 ^a (0.111)	0.195 ^a (0.094)	-0.216 ^a (0.063)	0.251 ^a (0.047)	-0.097 ^a (0.018)	-0.310 ^a (0.086)	0.261
Peru (1980.Q1 to 2005.Q3, N=94)							
(2,1,0)*	0.173 ^a (0.063)	0.142 ^a (0.158)				-0.024 ^b (0.013)	0.200
(2,1,1)	0.186 ^b (0.094)	0.139 ^a (0.065)			-0.058 (0.143)	-0.026 (0.016)	0.173
Poland (1985.Q4 to 2002.Q4, N=69)							
(2,1,0)*	0.474 ^a (0.016)	0.113 ^a (0.052)				-0.053 ^a (0.024)	0.215
(2,1,1)	0.476 ^a (0.010)	0.111 ^a (0.049)			-0.007 (0.045)	-0.053 ^a (0.024)	0.275
Uruguay (1985.Q1 to 2005.Q3, N=83)							
(2,1,0)	0.218 ^a (0.091)	0.290 ^a (0.116)				-0.063 ^a (0.029)	0.153
(2,1,1)*	0.265 ^b (0.147)	0.215 ^a (0.055)			-0.093 ^a (0.034)	-0.057 ^b (0.033)	0.196

NOTES: Maximum likelihood estimates. Figures in parentheses are robust (consistent) standard errors. ^{a, b} 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 test suggest uncorrelated and normally distributed residuals. The preferred specifications are with a *.

spanning roughly from the mid-1980s to the mid-2000s. 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 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 two lags are needed, but in Mexico when the lag length is four. 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 *.

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-90s, have dollarization ratios less than 30%. There are however, common features among all four 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,

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 analysed.

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 behaviour, $|a| < 1$. 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 behaviour 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 $A_3 - A_1A_2$ 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 centred on how fast each adjusts her forecast (and hence her portfolio) as news become available. We dubbed these 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

$$(13) \quad \begin{aligned} r_{it+1} &= r_{it} + \omega_{it+1} & \omega_{it} &\sim iid(0, \sigma_\omega^2) \\ S_t &= r_{it} + \epsilon_{it} & \epsilon_{it} &\sim iid(0, \sigma_{\epsilon_i}^2) \end{aligned}$$

and now we allow the variance of the noise to vary, $\sigma_{\epsilon_i}^2$. The signal-to-noise ratio is $q_i = \sigma_\omega^2 / \sigma_{\epsilon_i}^2$ 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 $\tilde{v}_{it} = v_{it} \sigma_{\epsilon_i}^{-2}$. The updating scheme of the MSE of \hat{r}_{it} , formerly equation (7), and its equilibrium value become:

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

³⁰ See Harvey (1989, ch. 4) and footnote 22.

whereas the forecasting rule, before, can be written now as:

$$(15) \quad \hat{r}_{it+1} = (1 - k_i) \hat{r}_{it} + k_i S_t, \quad \text{where } k_i = \frac{\tilde{v}_i}{\tilde{v}_i + 1}.$$

Given the optimal dollar investment follows the process:

$$(16) \quad x_{it} = (1 - k_i) x_{it-1} + \frac{1}{\tilde{v}_i + 1} S_t.$$

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.

5.2. Aggregation and further econometric issues

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

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

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:

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

$$(18) \quad \Delta X_t = \sum_{r=1}^{p_X} A_r \Delta X_{t-r} + \sum_{r=0}^{p_R} \beta_r \Delta R_{t-r} + U_t^*$$

where p_X and p_R are finite lag lengths. The presence of R , 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:

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

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 USD, respectively, δ_t the nominal depreciation (i.e., the percentage change of the nominal exchange rate, domestic currency per USD) and π_t the CPI inflation. We entertain two measurements of the real returns:

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

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.

³² Unfortunately we could not find time series long enough of country

Finally, the presence of a contemporaneous return 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

TABLE 2. AUGMENTED EQUATIONS

	A_1	A_2	A_3	$A_3 - A_1A_2$	$H_0: \beta = 0$	p_R	\bar{R}^2
Mexico (1985.Q4 to 2005.Q3, N=77)							
ex ante	0.291 ^a (0.096)	0.202 ^a (0.066)	-0.273 ^a (0.092)	-0.331 ^a (0.113)	11.50 ^a [0.000]	2	0.554
ex post	0.129 (0.108)	0.287 ^a (0.089)	-0.240 ^a (0.080)	-0.278 ^a (0.091)	22.56 ^a [0.000]	2	0.565
Peru (1980.Q1 to 2005.Q3, N=89)							
ex ante	0.242 ^a (0.043)	0.195 ^a (0.047)	0.003 (0.053)	-0.047 ^a (0.015)	9.086 ^a [0.000]	3	0.435
ex post	0.501 ^a (0.098)	0.138 ^b (0.083)	-0.027 (0.068)	-0.069 ^b (0.036)	30.85 ^a [0.000]	2	0.649
Poland (1985.Q4 to 2002.Q4, N=68)							
ex ante	0.449 ^a (0.043)	0.132 ^a (0.058)	-0.002 (0.049)	-0.059 ^a (0.022)	1.638 [0.203]	3	0.275
ex post	0.586 ^a (0.077)	0.164 ^a (0.070)	-0.123 (0.160)	-0.096 ^a (0.043)	2.402 ^b [0.099]	4	0.394
Uruguay (1985.Q1 to 2005.Q3, N=80)							
ex ante	0.252 ^a (0.104)	0.280 ^a (0.114)	-0.109 (0.0140)	-0.070 ^b (0.038)	3.153 ^a [0.049]	3	0.124
ex post	0.267 ^a (0.103)	0.349 ^a (0.117)	-0.073 (0.143)	-0.093 ^b (0.047)	2.189 [0.119]	2	0.152

NOTES: 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 error. ^a, ^b denotes significance at 5% [10%] level. Figures in the $H_0: \beta = 0$ column are F-statistics, p -values shown in braces. Lag length p_R was chosen to minimise 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.

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 USD, at various terms) instead of the US deposit rate.

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 minimisation 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 loses 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.

³³ We did not find a significant co-integration 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 years 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.

6. Concluding remarks

In countries with high dollarization ratios, participation in the dollar deposit market has become massive. Financial deregulation, liberalisation, innovation and informal currency markets have allowed a very heterogeneous 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 a 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.

Appendices

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,

$$(A1) \quad (1 - a_i L)x_{it} = a_i \left(\frac{1}{v_i + \sigma_\epsilon^2} \right) \xi_{it}, \quad \text{for } i = 1, 2.$$

After generating a common lag polynomial for both processes we have that:

$$(A2) \quad \begin{aligned} (1 - a_2 L)(1 - a_1 L)x_{1t} &= a_1 \left(\frac{1 - a_2 L}{v_1 + \sigma_\epsilon^2} \right) \xi_{1t} \\ (1 - a_1 L)(1 - a_2 L)x_{2t} &= a_2 \left(\frac{1 - a_1 L}{v_2 + \sigma_\epsilon^2} \right) \xi_{2t}. \end{aligned}$$

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:

$$(A3) \quad (1 - a_1 L)(1 - a_2 L)X_t = n_1 a_1 \left(\frac{1 - a_2 L}{v_1 + \sigma_\epsilon^2} \right) \xi_{1t} + n_2 a_2 \left(\frac{1 - a_1 L}{v_2 + \sigma_\epsilon^2} \right) \xi_{2t}.$$

Define $\tilde{\xi}_{it} + n_1 a_i (v_i + \sigma_\epsilon^2)^{-1} \xi_{it}$ for $i = 1, 2$. Then, (A3) boils down to:

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

We have that the aggregate dollarization ratio follows an ARMA(2,1) process. This simple example can be generalised 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 behaviour 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 $N \rightarrow \infty$, 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.

B. Aggregation

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

B.1 Equations (10) and (11)

Consider equation (9):

$$(B1) \quad x_t = ax_{t-1} + m + u_t,$$

where $m = u(1-a)v^{-1}$ and $u_t = a(v + \sigma_\epsilon^2)^{-1} \xi_t$ 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, $\mathcal{E}[z] = \int z dF_a$, such that $X_t = \mathcal{E}[x_t]$, $M = \mathcal{E}[m]$ and $U_t = \mathcal{E}[u_t]$. Aggregation of (B1) renders

$$(B2) \quad X_t = \mathcal{E}[ax_{t-1}] + M + U_t.$$

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:

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

where $\text{cov}(\alpha_s, m)$ is the cross-sectional covariance of α_s and m which is time-invariant. On the other side, $\text{cov}(\alpha_s, u_{t-s})$ 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 shows a recursion between $\mathcal{E}[\alpha_s x_{t-s}]$ and $\mathcal{E}[\alpha_{s+1} x_{t-(s+1)}]$. After solving it,

$$\text{(B4)} \quad \mathcal{E}(a x_{t-1}) = \sum_{r=1}^{\infty} A_r X_{t-r} + \sum_{r=1}^{\infty} \text{cov}(\alpha_r, m) + \sum_{r=1}^{\infty} \text{cov}(\alpha_r, u_{t-r}).$$

Let $V_t = \sum_{r=1}^{\infty} \text{cov}(\alpha_r, u_{t-r})$ and $\tilde{V} = E[V_t]$, where E is the expectation operator over time. Define also $\tilde{M} = M + \sum_{r=1}^{\infty} \text{cov}(\alpha_r, m) + \tilde{V}$ 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 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:

$$\text{(B5)} \quad \Delta x_t = a \Delta x_{t-1} + u_t - u_{t-1},$$

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),

$$\text{(B6)} \quad \mathcal{E}(a \Delta x_{t-1}) = \sum_{r=1}^{\infty} A_r \Delta X_{t-r} + V_t - V_{t-1},$$

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

so that ΔX_t can be written as:

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

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. 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:

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

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.

C. A brief note on fractional integration

Consider the univariate dynamic model:

$$(C1) \quad \Phi(L)(1-L)^d X_t = \Theta(L)\eta_t,$$

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 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. ESTIMATED FRACTIONAL INTEGRATION PARAMETER IN DOLLARIZATION RATIOS

	\hat{d}	$H_0: d = 0$		$H_0: d = 1$	
		<i>t-stat</i>	<i>p-value</i>	<i>t-stat</i>	<i>p-value</i>
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

NOTES: 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. 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.

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.

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

References

- Agénor, Pierre-Richard, and Nadeem U. Haque (1996), “Macroeconomic Management with Informal Financial Markets”, *International Journal of Finance and Economics*, vol. 1, n^o 2, pp. 87-101.
- Aoki, Kosuke (2003), “On the Optimal Monetary Policy Response to Noisy Indicators”, *Journal of Monetary Economics*, vol. 50, n^o 3, pp. 501-23.
- Baillie, Richard (1996), “Long Memory Processes and Fractional Integration in Econometrics”, *Journal of Econometrics*, vol. 73, n^o 1, pp. 5-59.
- Calvo, Guillermo A., and Carmen M. Reinhart (2002), “Fear of Floating”, *Quarterly Journal of Economics*, vol. 107, n^o 2, pp. 379-408.
- Céspedes, Luis F., Roberto Chang and Andrés Velasco (2004), “Balance Sheets and Exchange Rate Policy”, *American Economic Review*, vol. 94, n^o 4, pp. 1183-1193.
- Cook, David (2004), “Monetary Policy in Emerging Markets: Can Liability Dollarization Explain Contractionary Devaluations?”, *Journal of Monetary Economics*, vol. 51, n^o 6, pp. 1155-1181.
- De Nicoló, Gianni, Patrick Honohan and Alain Ize (2005), “Dollarization of Bank Deposits: Causes and Consequences”, *Journal of Banking and Finance*, vol. 29, n^o 7, pp. 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*, vol. 14, n^o 2, pp. 227-238.
- Guidotti, Pablo E., and Carlos A. Rodríguez (1992), “Dollarization in Latin America: Gresham’s Law in Reverse?”, *IMF Staff Papers*, vol. 39, n^o 3, pp. 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 Terence C. Mills and Kerry Patterson (eds.), *Palgrave Handbook of Econometrics*, Palgrave Macmillan.
- Ize, Alain, and Eduardo Levy Yeyati (2003), “Financial Dollarization”, *Journal of International Economics*, vol. 59, n^o 2, pp. 323-347.
- Kamin, Steven B., and Neil R. Ericsson (2003), “Dollarization in Post-Hyperinflationary Argentina”, *Journal of International Money and Finance*, vol. 22, n^o 2, pp. 185-211.
- Levy Yeyati, Eduardo (2006), “Financial Dollarization: Evaluating the Consequences”, *Economic Policy*, forthcoming.
- Lewbel, Arthur (1994), “Aggregation and Simple Dynamics”, *American Economic Review*, vol. 84, n^o 4, pp. 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*, vol. 51, n^o 7, pp. 1321-1352.
- Morón, Eduardo, and Diego Winkelried (2005), “Monetary Policy Rules for Financially Vulnerable Economies”, *Journal of Development Economics*, vol. 76, n^o 1, pp. 23-51.
- Pesaran, M. Hashem (2003), “Aggregation of Linear Dynamic Models: An Application to Life-Cycle Consumption Models under Habit Formation”, *Economic Modelling*, vol. 20, n^o 2, pp. 383-415.
- Quispe, Zenón (2000), “Monetary Policy in a Dollarized Economy: the Case of Peru”, in Lavan Mahadeva 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, n^o 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*, Edward Elgar, NH.
- Sturzenegger, Federico (1997), “Understanding the Welfare Implications of Currency Substitution”, *Journal of Economic Dynamics and Control*, vol. 21, n^{os} 2/3, pp. 391-416.
- Uribe, Martín (1997), “Hysteresis in a Simple Model of

- Currency Substitution”, *Journal of Monetary Economics*, vol. 40, n° 1, pp. 185-202.
- Zaffaroni, Paolo (2004), “Contemporaneous Aggregation of Linear Dynamic Models in Large Economies”, *Journal of Econometrics*, vol. 120, n° 1, pp. 75-102.

Índice

	<i>Pág.</i>
Presentación	<i>vii</i>
1. Motivación	1
2. Dos casos ilustrativos	9
2.1. Perú a inicios de la década del 2000	11
2.2. Polonia hacia una economía de mercado	14
2.3. Moraleja	15
3. Un modelo simple	17
3.1. Decisión de cartera	19
3.2. Pronósticos	20
3.3. Dinámicas individuales	23
3.4. Dinámica agregada	24
4. Evidencia empírica	27
4.1. Aspectos econométricos y datos	29
4.2. Resultados	31
5. Una interpretación	35
5.1. Reformulación del modelo	37
5.2. Agregación y aspectos econométricos adicionales	38
5.3. Resultados	41
6. Comentarios finales	43
Apéndices	47
A. La distribución de dotaciones y capacidades	49
B. Agregación	50
C. Una breve nota sobre integración financiera	52
Referencias	55

Index

	<i>Pág.</i>
1. Motivation	63
2. Two illustrative cases	71
2.1. Peru in the early 2000s	73
2.2. Poland towards a market economy	75
2.3. Moral	77
3. A simple model	79
3.1. Portfolio decision	81
3.2. Forecasting	82
3.3. Individual dynamics	84
3.4. Aggregate dynamics	85
4. Empirical evidence	89
4.1. Econometric issues and data	91
4.2. Results	93
5. A reinterpretation	95
5.1. Reformulating the model	97
5.2. Aggregation and further econometric issues	98
5.3. Results	101
6. Concluding remarks	103
Appendices	107
A. The distribution of endowments and abilities	109
B. Aggregation	110
C. A brief note on fractional integration	112
References	115

Este libro se terminó de imprimir durante
noviembre del 2009, en los talleres de
Alejandro Duplancher, Av. Coyoacán
nº 1012-2, México, D. F., 03100.
300 ejemplares.

CENTRO DE ESTUDIOS MONETARIOS LATINOAMERICANOS

Asociación Regional de Bancos Centrales

www.cemla.org

Estudios

ISBN: 978-607-7734-03-1



9 786077 734031