A Theory of Speculative Bubbles and Overshooting During Currency Crises

Antonio Doblas-Madrid*

Michigan State University
110 Marshall-Adams Hall
East Lansing, MI 48823
Phone: (517) 355 7755
Fax: (517) 432 1068
Email: doblasma@msu.edu

Abstract

In this paper, I propose bubbles à la Abreu and Brunnermeier (2003) as a theory that might explain some recent cases of exchange rate overshooting during currency crises. In my model, after a peg collapses, the domestic currency depreciates quickly. Bubbles arise because rational investors keep bidding up foreign currency even after learning that it is no longer undervalued. The exchange rate overshoots until the bubble bursts and the domestic currency regains some ground. The model has two main implications for currency crisis theory. First, self-fulfilling crises become more likely. Second, high interest rates can reduce, or even eliminate, bubbles.

Keywords: Currency Crises, Overshooting, Bubbles
JEL Codes: F31, F32, G12

* I would like to thank Timothy Kehoe, Patrick Kehoe, Juan Rubio-Ramirez, Leonardo Auernheimer, and participants at various seminars and conferences for helpful comments. All remaining errors are my own.

1. Introduction

During several recent currency crises, there was substantial exchange rate overshooting. As Figure 1 illustrates, in Belgium, Denmark, France and Ireland after the widening of ERM bands in August 1993, Korea and Thailand in 1997/1998, and Brazil in 1999, there was a period of fast depreciation, after which the currencies quickly regained a large fraction—in some instances, all—of the lost ground before stabilizing.

[Insert Figure 1 here]

In some episodes of overshooting, fundamentals plausibly explain the behavior of the exchange rate. For example, the Korean won lost over half of its value vis-à-vis the US dollar between October 1997 and January 1998 and begun recovering on January 28, 1998, exactly when, as reported by Blustein (2001), international bankers and Korean officials reached an agreement to roll over short-term debt owed by Korean banks. This reduced the severity of Korea's banking crisis, making it less likely that Korean authorities would resort to seignorage-financed bailouts. Within a couple of months, the won regained over half of the value lost in 1997, and stabilized. Brazil in 1999 is another overshooting case where fundamental events—again, narrated in detail in Blustein (2001)—easily explain the currency's recovery. Following sharp depreciation, the réal began recovering on March 4, coinciding with the arrival of a new central bank president, a bold interest rate hike, fiscal tightening and agreements with the IMF and private creditors.

In other overshooting episodes, however, the timing of fundamental events and the timing of the currency's recovery do not match. One such episode, also chronicled in detail by Blustein (2001), is Thailand in 1997/1998, which abandoned its peg (at roughly

25 baht/dollar) on July 2, 1997. While Thailand's first IMF program began in August 1997, the fundamental outlook did not really improve until November of that same year. A new government took office on November 9, 2007, the IMF approved a second program on November 25, and further loan disbursements were extended on December 8 as a reward for the new government's implementation of tough financial-sector reforms. (For Mr. Camdessus' exact statement, see IMF External Relations Department News Brief No. 97/29). But the price of a dollar continued skyrocketing from 41.5 baht on December 8 to 55.8 baht on January 12, 1998. At that point, in the absence of any important news about fundamentals, the baht began a quick recovery, reaching 37.8 baht/dollar by March 27, 1998 and ending the year at 36.3. Another crisis, described in detail by Buiter, Corsetti and Pesenti (1998), in which recovery is difficult to explain based on fundamentals is the widening of ERM bands in August 1993. The Belgian and French francs, Danish krona, and Irish punt quickly lost between 4 and 7 percent against the German mark.[1] Then, in the absence of important fundamental news, these currencies reversed course and regained most or all of the lost value by the end of the year.[2]

Motivated by these latter overshooting episodes, I propose a theory of bubbles in the foreign exchange market. Bubbles have traditionally been difficult to reconcile with standard economic theory, as they are often ruled out by backward induction.[3] Given a finite bursting time T, investors would sell at $T-1$, causing the bubble to burst at $T-1$. But then, investors would sell at $T-2$, and so on. Iterating, one concludes that prices

[1] In fact, France, Belgium/Luxembourg and Ireland joined the euro in 1999, with exchange rates vis-à-vis the German mark at the central parities of the narrow bands before August 1993. And the krona is currently pegged to the euro at the same central parity as before August 1993, with narrow $\pm 2.25\%$ bands.
[2] It shall be noted that four percent per month (week) corresponds to over 56 (700) percent per annum. Unless interest rates are very high, investors can make large profits if they anticipate this movement.
[3] Tirole (1982) and Santos and Woodford (1997) show that, in a wide variety of environments, bubbles are either fragile or inconsistent with equilibrium.

cannot deviate from fundamentals. Nevertheless, a series of recent, seminal papers have made great progress modeling bubbles in ways that are increasingly compatible with standard theory and that are immune to this backward-induction argument.[4] In particular, I will base my theory of bubbles on Abreu and Brunnermeier (2003) (AB henceforth). Roughly, the key idea in AB is that of a *greater fool's bubble*, by which it is optimal to invest in an overvalued asset, as long as there is a good chance of finding a *greater fool* who will pay even more later. Private information is crucial. Investors have different beliefs regarding when the crash will be, and do not know whether they expect the bubble to burst before or after others do. Investors understand that they will make profits if they sell before the crash and suffer losses if they end up being the *greater fool*, unable to unwind their position on time. Despite this risk, if probabilities and payoffs are such that expected profit is positive, they rationally choose to ride the bubble.[5]

The AB model is motivated by bubbles in stock markets, but can be extended to the currency crisis context. In AB, bubbles follow the arrival of new technologies, like the internet in the 1990s. At first, this fundamental shock drives the increases in asset prices in the industry. But a bubble arises because prices keep rising even after surpassing the level granted by fundamentals. The bubble grows for some time until it bursts and prices stabilize at a level higher than the initial level, but lower than the peak. In the currency overshooting context, there is an initial fixed parity at which foreign currency is undervalued vis-à-vis the domestic currency. The fundamental shock is the peg's collapse, after which the domestic/foreign exchange rate rises quickly as the mispricing

[4] An incomplete list of references includes Allen and Gorton (1993), Allen, Morris and Postlewaite (1993), Abreu and Brunnermeier (2003), Scheinkman and Xiong (2003), and Conlon (2004). See Brunnermeier (2001) for a survey.

[5] While details differ, Allen, Morris and Postlewaite (1993) and Conlon (2004) build on similar ideas. In fact, to my knowledge, it was Conlon (2004) who first coined the term *models of greater fool's bubbles*.

brought on by the fixed exchange rate is corrected. The exchange rate overshoots because investors, lured by large potential profits, keep betting against the domestic currency even after learning that it is no longer overvalued. They understand that the domestic currency will, at some point, reverse course and appreciate, but hope to continue making profits for some time while the bubble grows, and then unwind their positions before it bursts.[6]

I present a simplified, discrete-time version of AB, where the bubble depends on the following factors. First, the size of the bubble is an increasing function of the growth rate of the price, i.e., the depreciation rate of the domestic currency. Duration is also an increasing function of the fraction of investors who manage to sell before the crash. Finally, bubble duration is a decreasing function of the opportunity cost of funds, i.e., the domestic interest rate. In sum, the better the chances of being among the 'lucky' investors, the greater the rate of depreciation of the currency, and the lower the domestic interest rate, the larger the bubble.

The possibility of bubbles has two main implications for currency crisis theory. First, self-fulfilling crises are more likely in a world where bubbles are possible than in a world where prices always reflect fundamentals. The model I present in this paper does not preclude bubbles in the case, reminiscent of the 1993 widening of ERM bands, where initial domestic overvaluation is minimal or zero, and the post-crash price, rather than stabilizing at some intermediate level, returns to the pre-crisis level. The ERM crisis is often cited as an example a crisis that is best understood through the lens of second-generation models pioneered by Obstfeld (1986, 1994). However, a key ingredient in

[6] The AB model has also been used to interpret the run-up to the collapse of a fixed exchange rate. In Afonso (2007), asymmetric information makes it impossible for speculators to perfectly coordinate their attacks. Thus, some speculators sell earlier than others, and the peg collapses when total sales exhaust the Central Bank's reserves. While Afonso's model features continuous time, uncertainty about the time of collapse makes it possible for the exchange rate to jump discretely when the peg is abandoned.

these models is that monetary policy must become expansionary after a successful attack. If it did not, as Obstfeld (1986) shows, the domestic currency would *appreciate* when investors unwound their positions, making speculative profits negative, and keeping the attack from happening in the first place. This need for expansionary post-collapse policy is a feature of second-generation models that is at odds with the findings of Eichengreen, Rose, and Wyplosz (1995), who conclude that there is no evidence that monetary policy became more expansionary after speculative attacks during the European 1992/1993 crisis.[7] By contrast, the bubbles model is consistent with this finding. If the collapse of a peg is followed by a bubble, the currency still depreciates after the attack, at least temporarily. Consequently, investors may still attack even if fundamentals are such that the government does not find it optimal to run expansionary policies ex-post.

Finally, the fact that bubble duration is a decreasing function of the interest rate has implications for interest rate policy. In cases where the domestic currency is not initially overvalued, high interest rates can, by precluding bubbles, preclude self-fulfilling crises. And when there is initial overvaluation, the model suggests that the controversial IMF recommendation to increase interest rates after a peg collapses can avoid speculative overshooting and help stabilize the currency. These implications about post-collapse interest rates represent another departure from existing literature, which typically focuses on the role of interest rates in the periods prior to the abandonment of the fixed exchange

[7] Another view of the widening of ERM bands is that investors were genuinely surprised by policies. This view, however, may be seen as deviating from rational expectations, since out of six countries that widened bands, markets were only right about Portugal, and wrong about the other five. (Belgium, Denmark, France and Ireland were expected to run expansionary policies, and did not. Spain was not expected to expand monetary policy, but then did, and devalued again in 1995.) In any case, the possibility of mistakes is not incompatible with the emergence of bubbles. Investors may have initially believed that expansionary policies were going to be implemented, but as in the AB model, they may have decided to ride the bubble once they changed their mind about the government's intentions.

rate, rather than on post-collapse interest rates (see, for instance, Obstfeld (1994) and Lahiri and Végh (2003)).

The rest of the paper is organized as follows. In section 2, I describe the model. In section 3, I define equilibrium and derive the duration of the bubble as a function of the model's parameters. In section 4, I discuss comparative statics in the context of currency crises. In section 5, I conclude.

2. The Model

The model is a discrete-time version of AB, simplified and adapted to the foreign exchange context. Time is infinite with periods labeled $t = \ldots, 0, 1, \ldots$. There are two assets, a domestic asset, which pays a gross return $R > 1$ and a foreign asset, which pays no interest.[8] There is a Central Bank, which, while $t \leq 0$, maintains a fixed exchange rate. That is, for $t \leq 0$, the price one unit of the foreign asset is fixed at $s_t = 1$ units of the domestic asset. The Central Bank's foreign exchange reserves at the beginning of period *t* are given by F_t. The fixed exchange rate is abandoned if F_t falls below a threshold $\underline{F}$. I assume that $F_0 - 1 < \underline{F}$.

There is a unit measure of rational, risk-neutral investors with discount factor δ. For simplicity, I assume that $\delta = 1/R$. Rational investors begin period *t* holding d_t units of the domestic asset and f_t units of the foreign asset. Every period *t*, given d_t and f_t, they choose d_{t+1} and f_{t+1} subject to the budget constraint

$$\frac{d_{t+1}}{R} + s_t f_{t+1} = d_t + s_t f_t. \qquad (1)$$

[8] The assumption that the foreign asset pays no interest simplifies the analysis, but is not essential.

Short-sales constraints limit the extent to which portfolios can be leveraged. In particular, both d_t and f_t must be nonnegative.[9] Since $R > 1,$ while $t \leq 0,$ it is optimal to have all wealth invested in the domestic asset. Thus, normalizing time-0 wealth to one, investors' portfolio for $t \leq 0$ is given by $d_t = \delta^t$ and $f_t = 0.$

At time 0, investors launch a speculative attack in which they sell their entire holdings of the domestic asset to the Central Bank, setting $d_1 = 0$ and $f_1 = 1.$ Since $F_0 < \underline{F} - 1,$ the Central Bank abandons the peg. After the attack, the domestic currency depreciates quickly. Specifically, for $t \geq 1$ and until the bubble bursts, $s_t / s_{t-1} = G > R.$ Following AB, I assume that the fast rise in s_t is fueled by demand from behavioral traders who follow a technical "trending" or "momentum" rule. While the exchange rate s_t keeps rising, they keep betting against the domestic asset. But behavioral traders' absorption capacity is limited to $\kappa < 1.$ Once foreign assets in the amount of κ have been sold to them, they cannot keep bidding up the foreign asset.[10]

Depreciation of the domestic asset at the rate G is justified by fundamentals for some time. However, behavioral traders keep exerting upward pressure on s_t even after it surpasses the level granted by fundamentals. The model's ability to generate bubbles hinges on the fact that rational agents find it optimal not to sell their foreign assets for

[9] A lower bound other than zero could be assumed without qualitatively affecting results.

[10] The presence of behavioral agents simplifies the analysis and keeps the model close to AB, who justify this assumption by pointing out that proponents of the efficient markets hypothesis, such as Fama (1965), do not claim that all traders are rational in reality. Rather, they claim that bubbles cannot arise because rational investors, who have more resources than behavioral traders (since $1 > \kappa$), would arbitrate away any deviation of prices from fundamentals. Moreover, the presence of behavioral agents may not be essential, since other models of fool's bubbles, in particular Allen, Morris and Postlewaite (1993) and Conlon (2004), feature only rational actors. On the other hand, in these two papers, bubbles that are not robust to small changes in parameters. If one parameter is changed by a little bit while the others are held constant, bubbles disappear.

some time after learning that s_t has overshot the fundamental value. The first period in which foreign assets are overvalued is denoted by t_0, a random variable with probability function φ given by

$$\varphi(t_0) = e^{-\lambda t_0}\left(e^{\lambda} - 1\right) \qquad \text{for all } t_0 = 1, 2, \ldots, \tag{2}$$

with $\lambda > 0$. When t_0 is realized, it is not perfectly observable. Instead, every period from t_0 to $t_0 + N - 1$, a mass $1/N$ of rational investors observe a signal revealing that the exchange rate s_t has become higher than is justified by fundamentals. This divides the unit mass of rational investors into N different types, indexed by $n \in \{t_0, \ldots, t_0 + N - 1\}$. An investor knows when she observed her signal, but not when others did, i.e., she knows n, but not t_0. Conditional on n, an investor's distribution of t_0 is given by

$$\varphi(t_0 \mid n) = \begin{cases} \dfrac{e^{-\lambda t_0}}{e^{-\lambda(\max\{1, n-(N-1)\})} + \cdots + e^{-\lambda n}} & \text{if} \quad \max\{1, n-(N-1)\} \le t_0 \le n \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

The fact that investors do not know "where they stand in the line" is the crucial uncertainty that makes fool's bubbles possible.

The exchange rate s_t increases at the rate G until sales of the foreign asset by rational agents reach κ, at which point, the bubble bursts and the domestic asset appreciates back to where it was right before becoming undervalued.[11] For simplicity, I assume that $\kappa = K/N$, with K being a positive integer smaller than N, so that exactly

[11] I assume that the existence of a bubble has no effect on the fundamental value. While this need not be true in practice, specifying a more general post-crash price (as AB do in the stock-market context) does not change the model's main results. Also, if we examined the Brazilian and Korean episodes through the lens of the model, we would have to assume that at t_0, in addition to the revelation that the domestic currency is no longer overvalued, policies are implemented that increase the fundamental value of the domestic currency.

K types manage to sell before the crash.[12] Letting T be the first period after sales reach κ, the exchange rate s_T falls back to G^{t_0-1}, and stays there afterwards. Thus, the exchange rate process is given by

$$s_t = \begin{cases} 1 & \text{if } t \leq 0 \\ G^t & \text{if } 1 \leq t < T-1 \\ G^{t_0-1} & \text{if } \quad T \leq t. \end{cases}$$

The assumption that s_t continues to grow at the rate G until sales by rational investors reach κ implies that the exchange rate is, to some extent, insensitive to changes in supply and demand. To see this, suppose, for instance, that the optimal strategy for an investor of type n is to sell at time $n+\tau^*$, with $\tau^* \geq 0$. By assumption, during periods $t_0+\tau^*$ through $t_0+K-1+\tau^*$, the exchange rate continues to grow at the rate G as if sales had not begun even though some types are selling. Only at time $t_0+K+\tau^*$, when K types have sold, the bubble abruptly bursts. AB point out that this is problematic, since it implies that, for a number of periods, changes in supply and demand do not affect prices. However, AB also conjecture that bubbles may still arise if prices always reflected sales, but sales had a noisy component. In Doblas-Madrid (2008), I verify this conjecture by developing a version of AB featuring multidimensional uncertainty and prices that respond to changes in selling pressure at all times. The extended model avoids the 'invisibility of sales' problem and generates bubbles in equilibrium, but the analysis is considerably more complex, since strategies and beliefs depend, not only on an agent's signal, but also on the price history. Thus, this version of the model can be seen as a reduced-form version of Doblas-Madrid (2008).

[12] If K is not an integer, formulas become more complicated, but the main results do not change.

Finally, AB assume that the bubble may burst either for endogenous reasons, i.e., once κ shares are sold, or for exogenous reasons at time $t_0 + \overline{\tau}$, where $\overline{\tau}$ is an exogenous maximum bubble duration. I have excluded $\overline{\tau}$ from the description of the environment since I will only analyze the case in which there is an endogenous crash a finite number of periods after t_0. One could think of $\overline{\tau}$ as being a part of the environment, but not binding, because it exceeds the endogenous duration of the bubble.

3. Equilibrium

A type-*n* investor's strategy is a function $f_{t+1}(d_t, f_t \mid n)$, specifying next-period holdings of the foreign asset (and given (1), by implication, of the domestic asset) as a function of the current portfolio. A type-*n* investor's belief at time *t* is given by $\mu_t(t_0 \mid n)$, which is a probability distribution over values of t_0. In a Perfect Bayesian Nash Equilibrium (PBNE), the actions dictated by $f_{t+1}(d_t, f_t \mid n)$ must be optimal given $\mu_t(t_0 \mid n)$, and, in turn, $\mu_t(t_0 \mid n)$ must be consistent with the strategies $f_{t+1}(d_t, f_t \mid n)$.

To illustrate a simple bubble equilibrium, consider an example where type-*n* investors (with $n \geq K$) move all their wealth from the domestic to the foreign asset at time 0, and reverse the move as soon as either the bubble bursts or τ^* periods pass since observing the signal.[13] Specifically, the equilibrium strategy is given by

$$f_{t+1}(d_t, f_t) = \begin{cases} 0 & \text{if } t \leq 0 \text{ or } t \geq t^* \\ 1 & \text{if } 1 \leq t < t^*, \end{cases} \tag{4}$$

with $t^* = \min\{n + \tau^*, K + \tau^*, T\}$.

[13] AB call these strategies trigger strategies, since agents sell their entire holdings of the overvalued asset at once, and never buy it again. The strategies are also symmetric, since τ^* does not depend on *n*. AB show that, in their continuous-time model, which features a transaction cost and a maximum exogenous duration of the bubble, if there are no coordinating sunspots, the equilibrium where agents follow strategies given by (4) is unique.

If $t_0 \geq K$ and everybody follows these strategies, the first K types sell at times $t_0+\tau^*,\ldots,t_0+K-1+\tau^*$. At time $T=t_0+\tau^*+K$ these sales are revealed, the remaining types exit the market, and the bubble bursts. In the special case where $t_0<K,$ types who observe the signal before or at time K sell at $K+\tau^*$. Then, if $t_0=1,$ the bubble bursts at time $K+\tau^*+1$. Otherwise, the remaining types $K+1,\ldots,t_0+K-1$ sell, respectively, at times $K+1+\tau^*,\ldots,t_0+K-1+\tau^*,$ and the crash happens at time $T=t_0+\tau^*+K.$

Having established that, if everybody follows (4), the bursting time is $t_0+K+\tau^*,$ we can describe $\mu_t(t_0\,|\,n)$, the equilibrium belief of a type-n investor at time t. Let us first assume that $t \geq n,$ i.e., let us first consider the beliefs of investors after observing the signal. The signal n reveals to the investor that t_0 is between $\max\{1, n-(N-1)\}$ and n. In turn, this implies that the bubble will burst no sooner than $\max\{1,n-(N-1)\}+\tau^*+K$ and no later than $n+\tau^*+K.$ Thus, while $t<\max\{1,n-(N-1)\}+\tau^*+K,$ the fact that the bubble has not burst reveals nothing about t_0 and $\mu_t(t_0\,|\,n)$ is the same as $\varphi(t_0\,|\,n)$, which is given by (3). On the other hand, if the bubble is still growing at time t, with $\max\{1,n-\left(N-1\right)\}+K+\tau^*\leq t<n+K+\tau^*,$ investors learn that $t_0>t-(K+\tau^*),$ and thus $\mu_t(t_0\,|\,n)$ becomes

$$\mu_t(t_0\,|\,n)=\begin{cases} \dfrac{1-e^{-\lambda}}{1-e^{-\lambda\{n+K+\tau^*-t\}}}e^{-\lambda\{t_0-[t+1-(K+\tau^*)]\}} & \text{if } t_0\in\left\{t+1-\left(K+\tau^*\right),\ldots,n\right\} \\ 0 & \text{otherwise.}\end{cases} \tag{5}$$

Finally, if $t<n,$ i.e., before observing the signal, agents only know that $t_0>t-(N-1)$, since they have seen no signal yet, and that $t_0>t-(K+\tau^*)$, since there has been no crash as of time t. Therefore, if $t<n,$

$$\mu_t(t_0 \mid n) = \begin{cases} e^{\lambda\xi}\varphi(t_0) & \text{if } t_0 > \xi \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where $\xi \equiv \max\{0, t-(N-1), t-(K+\tau^*)\}$.

Having described equilibrium beliefs, I next find conditions under which each agent agrees to follow the equilibrium strategy, and use those conditions to derive the equilibrium value of τ^*.

PROPOSITION 1 *If* $\pi = (1-e^{-\lambda})/(1-e^{-\lambda K})$, $(1-\pi)G < R < \pi G^{-K} + (1-\pi)G$, *and* $G > e^{\lambda}$, *there exists a PBNE with* $0 < \tau^* < \infty$, *in which agents play strategies given by* (4) *and bubble duration* τ^* *equals* $\left\lceil -K - \left\{\ln\left(R-\left(1-\pi\right)G\right) - \ln\pi\right\}/\ln G \right\rceil$, *where the ceiling function* $\lceil x \rceil$ *rounds* x *up to the nearest integer greater or equal than* x.

PROOF Since $R > 1$, investors hold all their wealth in domestic assets before the peg's collapse and after the burst of the bubble. That is, since $R > 1$, investors set $f_t = 0$ for $t \leq 0$ and $t > T$. Similarly, investors set $d_1 = 0$ and $f_1 = 1$ at time 0 because $G > R$.

After time 0, type-*n* investors must be willing to hold foreign assets until the bubble bursts or t equals $n + \tau^*$. And if the bubble does not burst after they sell, they must be willing to continue holding domestic assets. To see under what conditions this is optimal, consider the sell-or-wait trade-off of a type-*n* investor who is fully invested in foreign assets at time *t*, with $t \geq n$. If she sells at *t*, she will obtain G^t and earn the domestic return *R* between periods t and $t+1$. If she waits, at time $t+1$ her foreign assets will be worth $G^{t-(K+\tau^*)}$ or G^{t+1}, depending on whether the bubble bursts at $t+1$ or not. Since the probability of a crash at $t+1$ given time-*t* information equals $\mu_t(t+1-(K+\tau^*) \mid n)$, selling foreign assets at t is optimal if

$$R \geq \mu_t(t+1-(K+\tau^*) \mid n)G^{-(K+\tau^*)} + [1-\mu_t(t+1-(K+\tau^*) \mid n)]G. \tag{7}$$

While $t+1 < \max\{1, n-(N-1)\} + K + \tau^*,$ the probability of a crash is zero, and thus, waiting is clearly optimal. Once $t+1$ reaches $\max\{1, n-(N-1)\} + K + \tau^*,$ the probability of a crash at $t+1$ given time-*t* information begins to be positive. Specifically, for $t+1$ between $\max\{1, n-(N-1)\} + K + \tau^*$ and $n + K + \tau^*,$ $\mu_t(t+1-(K+\tau^*)\,|\,n)$, as given by (5), equals $(1-e^{-\lambda})/(1-e^{-\lambda\{n+K+\tau^*-t\}}).$

In equilibrium, investors with $n \geq K$ are supposed to sell at time $t = n + \tau^*.$ Given (5), this means that $\mu_{n+\tau^*}(n+1-K\,|\,n)$ equals $(1-e^{-\lambda})/(1-e^{-\lambda K})$. Letting

$$\pi = \frac{1-e^{-\lambda}}{1-e^{-\lambda K}}, \tag{8}$$

we can write (7) for $t = n + \tau^*$ as

$$R \geq \pi G^{-(K+\tau^*)} + (1-\pi)G. \tag{9}$$

If (9) holds, type-*n* investors are willing to sell their foreign assets at time $n+\tau^*.$ Furthermore, they want to continue holding domestic assets even if the bubble does not burst after they sell. To see why, note that, at $t > n + \tau^*,$ domestic assets yields *R*, while the expected return on foreign assets is the right-hand side of (7), with $\mu_t(t+1-(K+\tau^*)\,|\,n) > \pi.$ In other words, if (9) holds, domestic assets have higher expected returns than foreign assets in period $t = n + \tau^*,$ and since the probability of a crash at $t+1$ is increasing in *t*, this is even more strongly the case after period $n+\tau^*.$

Equilibrium also requires that, unless the bubble bursts, investors do not sell their foreign assets before period $n+\tau^*.$ In a continuous-time model, such as the original AB model, τ^* would make (9) hold with equality. In that case, type-*n* investors would not sell at time $t = n + \tau^* - j$ for any $j > 1,$ since the probability of a crash at $n+\tau^*-j+1$

would be $(1-e^{-\lambda})/(1-e^{-\lambda(K+j)})$, less than π. While this idea, in broad terms, also applies in this model, discrete time requires τ^* to be an integer. Therefore,

$$\tau^* = \left\lceil -K - \frac{\ln\left(R-(1-\pi)G\right)-\ln\pi}{\ln G} \right\rceil. \tag{10}$$

That is, the equilibrium τ^* is the smallest integer greater or equal than the value that makes (9) hold with equality. This upwards rounding increases the value of τ^* and hence the size of the crash. This might conceivably be problematic, since it could make it optimal for investors of type $n+1$ to sell at time $n+\tau^*$. To address this concern, in appendix A, I prove that $G > e^{\lambda}$ is sufficient to ensure that this never happens.[14]

To see under what parameter restrictions τ^* is finite and positive, note that, if $R \le (1-\pi)G$, type-*n* agents do not want to sell at any time $n+\tau^*$, since no matter how high τ^* is, (9) can never hold. On the other end, if $R \ge \pi G^{-K} + (1-\pi)G$, type-*n* investors sell at time *n* and no bubble arises.

Investors with $n < K$, if they exist, do not to sell at time $n+\tau^*$, since they know that the bubble cannot possibly burst before period $1+K+\tau^*$, and thus have absolutely no incentive to sell before period $K+\tau^*$. At time $K+\tau^*$, however, they definitely want to sell, since their sell-or-wait trade-off is governed by a modified version of (9), with the probability of a crash given by $(1-e^{-\lambda})/(1-e^{-\lambda n}) > \pi$.

Finally, type-*n* investors do not want to sell when $t < n$, i.e., before observing the signal. From the point of view of these agents, the probability of a crash at $t+1$ is

[14] While, intuitively, I suspect that, unless $G > e^{\lambda}$, it is impossible to have $R < \pi G^{-K} + (1-\pi)G$, I have not been able to prove this. For this reason, $G > e^{\lambda}$, appears as a separate assumption in this proposition.

$(1-e^{-\lambda})$, which is clearly below $(1-e^{-\lambda})/(1-e^{-\lambda(K+1)})$. Since it is optimal for type-n investors not to sell at $n+\tau^*-1$, it is also optimal for them not to sell while $t<n$. **Q.E.D.**

4. Comparative Statics in the Context of Currency Crises

While the equilibrium level of τ^* is explicitly given by (10), it is actually easier to understand the comparative statics of the model by examining inequality (9). Clearly, the equilibrium bubble duration τ^* is always a decreasing function of R and π. On the other hand, the relationship between τ^* and the rate of growth of the bubble G is non-monotonic. For small values of G, it is possible that τ^* falls as G increases. But as G continues to increase, τ^* becomes an increasing function of G. In fact, as G approaches $R/(1-\pi)$, τ^* approaches infinity. Also, the probability π is derived from the more primitive parameters λ and K. Increases in λ always reduce bubble duration, since π is strictly and monotonically increasing in λ. Finally, K, which is defined as κN, may, like G, have a non-monotonic effect on τ^*. For low values of K, it is possible that τ^* decreases with K. But past a certain point, the net effect of an increase in K on τ^* is positive, since increases in K lead to declines in π.

In the context of currency crises, the parameters G and R have obvious empirical counterparts. The growth rate of the exchange rate G can be directly observed, and is typically large, since exchange rates move very quickly after speculative attacks dismantle currency pegs. The parameter R represents the domestic interest rate, which is also observable. In actual crisis episodes, R is higher than in tranquil times, but it is still, in most cases, far below G. As mentioned in the introduction, the fact that the duration of the bubble is inversely related to the domestic interest rate suggests a rationale, not only for high pre-collapse domestic interest rates, but also for high post-collapse interest rates in order to stabilize the currency, and limit the degree of overshooting.

Parameters K and λ do not have such obvious empirical counterparts. Nevertheless, it is possible to get a good sense of the role that they might have played in different crisis episodes. We shall recall that *K* is defined as $\kappa N,$ where *N* is the number of types in the economy and κ the fraction of foreign assets that can be sold before the bubble bursts. The number of types *N* is a measure of the degree of uncertainty regarding $t_0.$ For instance, in the cases of Brazil and Korea, there was a clear date at which policies changed and the fundamental outlook improved. From the point of view of the model, *N* would be very small, possibly 1, in these cases, and thus, no bubble would arise. On the other hand, in Thailand, it is much more difficult to pinpoint when t_0 occurred. Was it November 9? November 25? December 8? Similarly, after the widening of ERM bands in 1993, one cannot point to a date at which it became clear that France, Ireland, Belgium or Denmark were not going to follow expansionary policies. The absence of such a clear date translates into a high value of *N*, and this ambiguity makes bubbles possible. To interpret κ, the key idea is that markets need to be noisy enough to hide the sales of some number of types. In this regard, the foreign exchange market is similar to the stock market, in the sense that they are both very volatile. Since exchange rates are remarkably noisy, it is very plausible that sales by some rational types could be mistaken for random day-to-day fluctuations.

Finally, it remains to discuss the role of $\lambda.$ This parameter has two effects. First, it determines the expected value of $t_0.$ Specifically, the higher $\lambda,$ the lower the expected $t_0.$ Also, increases in λ lead to increases in $\pi.$ When interpreting crisis episodes through the lens of the model, λ would be low (and, consequently, the expected t_0 would be high) in cases, such as Brazil, Thailand, or Korea, where the initial fixed exchange rate was substantially overvalued. On the other hand, λ would be high (and, consequently, the expected t_0 would be low) in cases, such as the widening of ERM bands in 1993, where the initial overvaluation was insignificant. In these latter scenario, it bears emphasizing that, even if a high value of λ means that π also tends to be high, if G and K are high enough, and R is low enough, τ^* can still be positive. This has important implications for self-fulfilling crises. To see why, imagine that λ was high, so that the fixed exchange rate was not significantly overvalued, and τ^* was zero. In that case, a small transaction cost at time 0 might dissuade investors from attacking the peg, and the fixed exchange rate might survive. On the other hand, if bubbles are possible, i.e. if τ^* is positive, the same transaction cost at time 0 might not be enough to preclude the attack, and a self-fulfilling attack might dismantle a sustainable peg, even if fundamentals are such that the Central Bank does not follow expansionary policies after it abandons the initial fixed exchange rate.

5. Conclusion

In this paper, I propose a theory of exchange rate overshooting after currency crises based on bubbles à la Abreu and Brunnermeier (2003). In my model, after an overvalued peg collapses, the domestic currency depreciates quickly. A bubble arises because rational investors keep shorting domestic currency even after learning that it is no longer overvalued. Foreign currency becomes temporarily overvalued until the bubble bursts and the currency returns to a level consistent with fundamental value. The possibility of bubbles has two main implications for theories of balance-of-payments crises. First, self-fulfilling crises become more likely, since even if the absence of post-collapse expansionary policy, the bubble causes the domestic currency to depreciate after the peg is abandoned, at least temporarily. Second, the model provides a rationale for high interest rates after the peg is abandoned, since high interest can reduce, or even eliminate, overshooting due to bubbles.

While models of greater fool's bubbles, such as the version of AB presented in this paper, have greatly advanced our understanding of bubbles, some of their features are still markedly unrealistic. For instance, in the models, bubbles burst abruptly, while in reality, bubbles often deflate gradually. However, making the model more realistic along this dimension would actually make it easier to generate bubbles, as it would greatly reduce the threat of a great sudden loss, hence increasing the expected profit for investors. Another limitation of the model concerns the fact that the model is silent about the relationship between prices and trading volume, whereas, in reality, trading volume tends to be abnormally high in periods where prices are booming or crashing.

For future work, it would be interesting to derive the implications of theories of greater fool's bubbles for the prices of derivatives and compare the implied prices to those actually observed. While forward exchange rates are pinned down by covered interest parity and may thus not be suitable for this purpose, it may be possible to test the empirical validity of models of fool's bubbles using options prices during episodes of exchange rate overshooting.

References

1. Allen, Franklin, and Gary Gorton. (1993). "Churning Bubbles." The Review of Economic Studies 60, 813-836.
2. Allen, Franklin, Stephen Morris and Andrew Postlewaite. (1993). "Finite Bubbles with Short Sales Constraints and Asymmetric Information." Journal of Economic Theory 61, 209-229.
3. Abreu, Dilip and Markus Brunnermeier. (2003). "Bubbles and Crashes." Econometrica 71, 173-204.
4. Afonso, Gara M., (2007). "Imperfect Common Knowledge in First-Generation Models of Currency Crises." International Journal of Central Banking 3(1), 81-112.
5. Blustein, Paul. (2001). The Chastening: Inside the Crisis that Rocked the Global Financial System and Humbled the IMF. New York: Public Affairs Books.
6. Brunnermeier, Markus. (2001) Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis, and Herding, Oxford University Press (2001).
7. Buiter, Willem, Giancarlo Corsetti and Paolo Pesenti. (1998). Financial Markets and European Monetary Cooperation. The Lessons of the 1992-93 Exchange Rate Mechanism Crisis, Cambridge University Press.
8. Conlon, John. (2004). Simple Finite Horizon Bubbles Robust to Higher Order Knowledge, Econometrica 72, 927-936.
9. Doblas-Madrid, Antonio. (2008). "A Robust Model of Bubbles with Multidimensional Uncertainty." Manuscript, Michigan State University.
10. Eichengreen, Barry, Andrew Rose and Charles Wyplosz. (1995). "Speculative Attacks on Pegged Exchange Rates: An Empirical Exploration with Special Reference to the European Monetary System." Working Papers in Applied Economic Theory 95-04, Federal Reserve Bank of San Francisco
11. Fama, Eugene. (1965). "The Behavior of Stock-Market Prices." Journal of Business, 38, 34-105.
12. International Monetary Fund External Relations Department, 1997, News Brief No. 97/29. IMF Executive Board Completes Review of Thailand Program, http://www.imf.org/external/np/sec/nb/1997/nb9729.htm, retrieved July 22, 2008.
13. Lahiri, Amartya and Carlos A. Végh. (2003). "Delaying the Inevitable: Interest Rate Defense and Balance of Payments Crises." Journal of Political Economy 111, 404-424.

14. Obstfeld, Maurice. (1986). “Rational and Self-Fulfilling Balance-of-Payments Crises.” The American Economic Review 76, 72–81.
15. Obstfeld, Maurice. (1994). “The Logic of Currency Crises.” Banque de France – Cahiers économiques et monétaires 43, 189–213.
16. Pacific Exchange Rate Service http://fx.sauder.ubc.ca/
17. Santos Manuel S. and Michael Woodford. (1997.) “Rational Asset Pricing Bubbles.” Econometrica 65, 19-57.
18. Scheinkman, Jose, and Wei Xiong. (2003.) “Overconfidence and Speculative Bubbles.” Journal of Political Economy 111, 1183-1219.
19. Tirole, Jean. (1982). “On the Possibility of Speculation under Rational Expectations.” Econometrica, 50, 1163-1182.

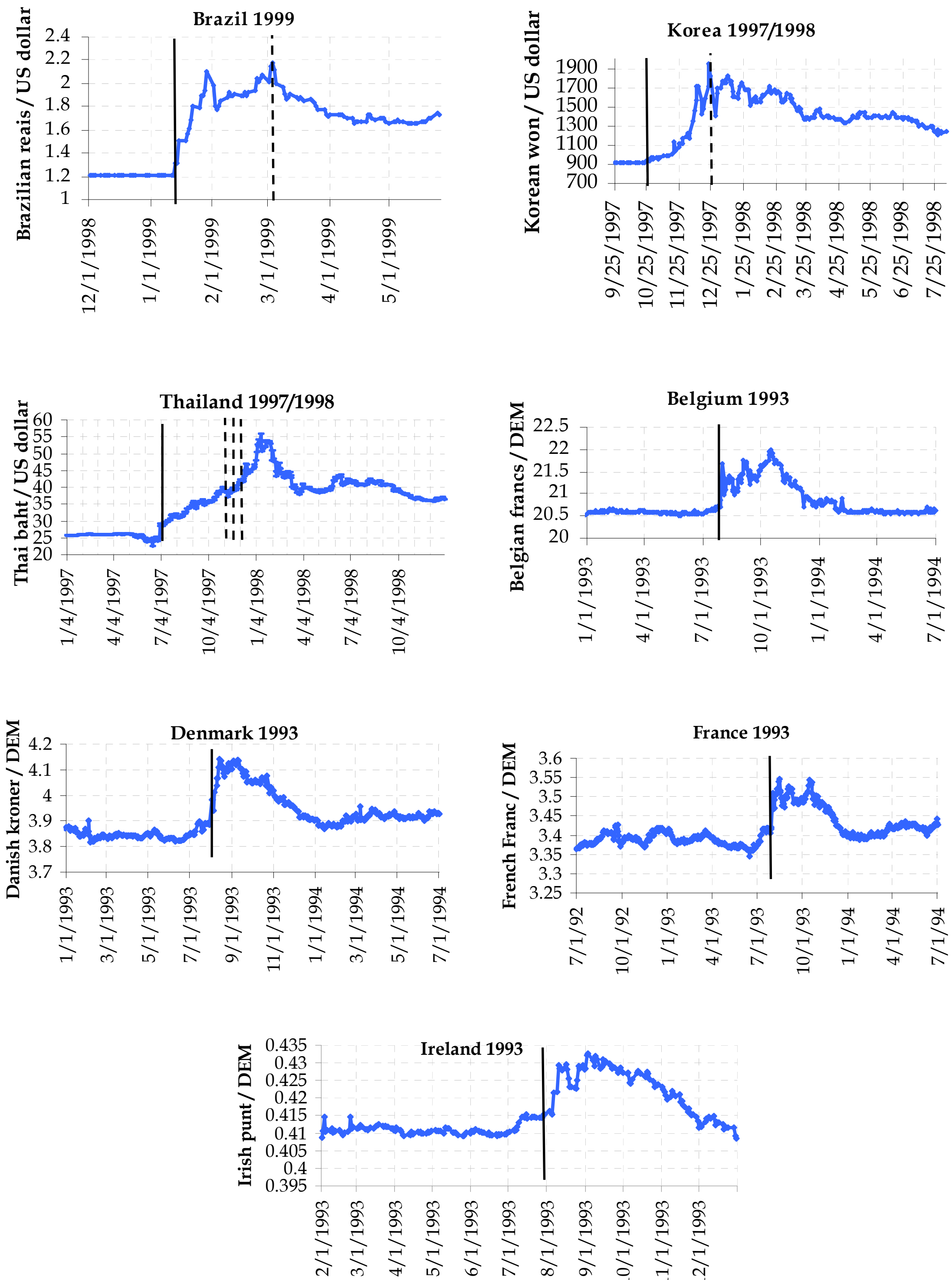

Figure 1 - Overshooting episodes. Solid lines show when pegs were abandoned, dashed lines when important fundamental changes became known. In Korea and Brazil, the dashed line marks the beginning of recovery. In Thailand, the dashed line could be placed at several dates. In Europe, there is even more ambiguity regarding when it became known that policy would not be expansionary. Source: Pacific Exchange Rate Service.

Appendix A – Proof that, if $G > e^{\lambda}$, investors of type $n+1$ do not sell at time $n+\tau*$

Since rounding up makes (9) hold with strict inequality, it is necessary to verify that investors of type $n+1$ do not wish to sell at time $t = n+\tau*$. For agents of type $n+1$, given the information available at time $t = n+\tau*$, the probability of a crash at time $t+1 = n+\tau*+1$ is given by

$$\mu_{n+\tau*}(n+\tau*+1 \mid n+1) = \frac{1-e^{-\lambda}}{1-e^{-\lambda((n+1)+K+\tau*-(n+\tau*))}} = \frac{1-e^{-\lambda}}{1-e^{-\lambda(K+1)}}.$$

Thus, investors of type $n+1$ wish to wait at time $n+\tau*$ as long as:

$$R < \frac{1-e^{-\lambda}}{1-e^{-\lambda(K+1)}} G^{-(\tau*+K)} + \left(1 - \frac{1-e^{-\lambda}}{1-e^{-\lambda(K+1)}}\right) G. \tag{A1}$$

Since $\tau*$ is the lowest integer greater than $-K - \left\{\ln\left(R-(1-\pi)G\right) - \ln\pi\right\}/\ln G$, this implies that type-n agents would not want to sell at time $n+\tau**$ if $\tau**$ was equal to $\tau*-1$. Therefore,

$$R < \frac{1-e^{-\lambda}}{1-e^{-\lambda K}} G^{-(K+\tau*-1)} + \left(1 - \frac{1-e^{-\lambda}}{1-e^{-\lambda K}}\right) G. \tag{A2}$$

Clearly, a sufficient condition for (A1) to hold is that the right-hand-side of (A1) be no smaller than that of (A2). That is,

$$\frac{1-e^{-\lambda}}{1-e^{-\lambda(K+1)}} G^{-(\tau*+K)} + \left(1 - \frac{1-e^{-\lambda}}{1-e^{-\lambda(K+1)}}\right) G \geq \frac{1-e^{-\lambda}}{1-e^{-\lambda K}} G^{-(K+\tau*-1)} + \left(1 - \frac{1-e^{-\lambda}}{1-e^{-\lambda K}}\right) G. \tag{A3}$$

Rearranging terms, this inequality becomes

$$\frac{1-e^{-\lambda}}{1-e^{-\lambda(K+1)}} \left(G^{-(\tau*+K)} - G\right) \geq \frac{1-e^{-\lambda}}{1-e^{-\lambda K}} \left(G^{-(K+\tau*-1)} - G\right)$$

$$\Leftrightarrow \frac{G^{-(\tau*+K)} - G}{1-e^{-\lambda(K+1)}} \geq \frac{G^{-(K+\tau*-1)} - G}{1-e^{-\lambda K}}.$$

Next, multiply both sides by $(1-e^{-\lambda(K+1)})$ and divide both sides by $(G^{-(K+\tau*-1)} - G)$. Since $(G^{-(K+\tau*-1)} - G)$ negative, the direction of the inequality is reversed, and we get

$$\frac{G^{-(K+\tau*)} - G}{G^{-(K+\tau*-1)} - G} \leq \frac{1-e^{-\lambda(K+1)}}{1-e^{-\lambda K}}.$$

Next, transform the left-hand side as follows,

$$\frac{G^{-(K+\tau*)} - G}{G^{-(K+\tau*-1)} - G} = \frac{G - G^{-(K+\tau*)}}{G - G^{-(K+\tau*-1)}} = \frac{1 - G^{-(K+\tau*+1)}}{1 - G^{-(K+\tau*)}},$$

so that the inequality becomes

$$\frac{1 - G^{-(K+\tau*+1)}}{1 - G^{-(K+\tau*)}} \leq \frac{1-e^{-\lambda(K+1)}}{1-e^{-\lambda K}}.$$

Since the left-hand side is an increasing function of $\tau*$, if this inequality holds for $\tau* = 0$, it will also hold for positive $\tau*$. Thus, a sufficient condition for (A3) to hold is

$$\frac{1 - G^{-(K+1)}}{1 - G^{-K}} \leq \frac{1-e^{-\lambda(K+1)}}{1-e^{-\lambda K}}. \tag{A4}$$

Note that both sides of this inequality are of the form $(1-a^{-(K+1)})/(1-a^{-K})$. This fraction can be rewritten as follows,

$$\frac{1-a^{-(K+1)}}{1-a^{-K}}=\frac{1-a^{-K}}{1-a^{-K}}+\frac{a^{-K}-a^{-(K+1)}}{1-a^{-K}}=1+\frac{a-1}{a^{K+1}-a}=1+\frac{1}{a}\frac{a-1}{a^{K}-1}=1+\frac{1}{a}\frac{1}{(1+a+a^{2}+\cdots+a^{K-1})}.$$

From here, it is clear that, for any $K\geq 1$, the higher a, the lower $(1-a^{-(K+1)})/(1-a^{-K})$. An immediate consequence of this is that (A4) holds if $G\geq e^{\lambda}$. **Q.E.D.**