A Framework for Studying the Monetary and Fiscal History of Latin America, 1960-2017

Timothy J. Kehoe, Juan Pablo Nicolini, and Thomas J. Sargent
MARCH 2019
A Framework for Studying the Monetary and Fiscal History of Latin America, 1960-2017*

Timothy J. Kehoe  
University of Minnesota,  
Federal Reserve Bank of Minneapolis,  
and National Bureau of Economic Research

Juan Pablo Nicolini  
Federal Reserve Bank of Minneapolis and Universidad Torcuato Di Tella

Thomas J. Sargent  
New York University and Hoover Institution

ABSTRACT

In this chapter, we develop a conceptual framework for analyzing the interactions between aggregate fiscal policy and monetary policy. The framework draws on existing models that analyze sovereign debt crises and balance-of-payments crises. We intend this framework as a guide for analyzing the monetary and fiscal history of a set of major Latin American countries from the 1960s until now.

Keywords: Monetary policy, Fiscal policy, Balance-of-payments crisis, Debt crisis.  
JEL Codes: E52, E63, F31, H63, N16.

* This paper is part of a book entitled The Monetary and Fiscal History of Latin America, funded by the Becker Friedman Institute at the University of Chicago. We thank Fernando Alvarez and Rody Manuelli for very useful comments on a previous draft and to the authors of the country studies for feedback at many stages of the project. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.
1. Introduction

It has been almost half a century since Eduardo Galeano published the first edition of *The Open Veins of Latin America as Las venas abiertas de América Latina* in 1971. Since then, more than seventy-five editions have been launched, and the book has been translated into more than a dozen languages. Galeano's book has a major virtue: it is an intellectual project that provides a diagnosis of a painful Latin American reality and attempts to build an explanation for the region's underperformance. In addition, it was the first such attempt to become part of the popular culture: the term *Open Veins*, or *Las venas abiertas* in Spanish, went beyond the limits of its readers, its argument, and its ideology, and came to occupy its own place in songs and newspaper articles. It became a popular icon all across the region, representing a generalized sense of failure.

Galeano’s book is the product of a particular time, and it provides an explanation for Latin America’s failure that we do not share. Nevertheless, Galeano's diagnosis of the underperformance of the region is uncontroversial: the great homeland that Simón Bolívar imagined at the dawn of the nineteenth century has been a profound disappointment for José Artigas, Miguel Hidalgo, Bernardo O'Higgins, José de San Martin, Antonio Sucre, Túpac Amaru, and many others who devoted their lives to independence, hoping for freedom and prosperity for the region. Two hundred years later, Latin America continues to be a region with very high income inequality and low social mobility. Only sub-Saharan Africa surpasses Latin America in terms of economic stagnation.

A similar view on the performance of the region can be found in the writings of Mario Vargas Llosa. His famous novel *Conversation in the Cathedral*, published in 1969 as *Conversación en la catedral*, immortalized the second sentence in its opening paragraph: “At what precise moment had Peru [messed] itself up?” (or in Spanish, “*En qué momento se había jodido el Perú*?”). That sentence summarizes a view of decadence that Vargas Llosa described for Peru in many of his novels, and for all of Latin America in his later writings, after he had moved into politics. Regarding the reasons for the decadence, Vargas Llosa has very different views from the ones expressed in *Las venas abiertas*, but the general sense of disappointment is common to both writers.
Good and comparable data across the region for the first century following independence are not available. Data for the twentieth century, however, can be used to illustrate the economic stagnation emphasized by Galeano and Vargas Llosa. We chose as a starting point the year 1935, so as to leave out the First World War and the start of the Great Depression, which were unusual events for the world as a whole.

We describe the data in more detail below, but as an introductory note, during the period from 1935 to 1973, when Vargas Llosa and Galeano had just published their books, the region had enormous difficulty bridging the income gap with rich countries. For example, the average income of the region grew from 21 percent relative to the United States in 1935 to 26 percent in 1973—a growth of only 5 percent in thirty-eight years. (We focus on the ten largest countries in South America plus Mexico.) In comparison, we note that in the same period, average income in the twelve countries of Western Europe—the EU 12, the twelve original members of the European Union—went from 58 percent to 76 percent relative to the United States—a growth of 18 percent—with a terrible war in the middle of that period.

These numbers represent averages for the region. If one cares about the most vulnerable groups in society, the situation is even worse, given that levels of economic inequality are greater in Latin America than in the United States. Therefore, the differences between the poor in Latin America and the poor in the United States are substantially greater than those mentioned above. The most evident symptom of this situation is the systematic migratory flow of workers from the south to the north. The bitter consolation for societies that have failed to generate opportunities for the most vulnerable is that many of them have successfully managed to find those opportunities in different societies.

This reality takes on a more dramatic dimension if we review the years from the publication of Las venas and Conversación to the end of the century. Data that are comparable to those discussed above reveal a strong deterioration relative to the United States in the final three decades of the twentieth century. On average, the region declined from the 26 percent that it had reached in 1973 back to 23 percent by 2000. It took the region the first decade and a half of the twenty-first century to bring its average back to the 26 percent it had reached in 1973.
The obvious and immediate question that arises is, Why? What went wrong in Latin America? The only honest answer is that we do not know. As a profession, we do not have the policy answers that would have guaranteed convergence of Latin America to the income level of the richest countries in the world. Coincidental with the periods of poor economic performance, however, countries in Latin America have been plagued by economic crises. The specific symptoms of each crisis have been very different: high inflation rates, balance of payments crises followed by large devaluations, banking crises, defaults on government debt, deposit confiscations, and so on.

Our fundamental hypothesis is that, despite their different manifestations, all economic crises in Latin America have been the result of poorly designed or poorly implemented macro-fiscal policies. The prototypical scenario for a crisis is that, because of social pressures, the government increased expenditures without a compensating increase in revenues. Initially, it financed the resulting deficit by borrowing, with a large fraction of the borrowing done abroad. When the debt reached a certain level, lenders were unwilling to lend more and a crisis unfolded. This process continued until there was a reform. In a number of countries, the process leading to a crisis occurred more than once. A reasonable conjecture is that the prevalence of crises is at the root of a sizable fraction of the stagnation of Latin America.

The first aim of the series of chapters in this book is to collect systematic and comparable data on several macroeconomic variables for the eleven countries included in this volume. We believe these variables are key to understanding the main causes of the sequence of crises that prevailed in the region. The second aim is to use this data set to construct narratives for each country, so the crises and the evolution of the main variables can be jointly understood within the economic environment of the time and the macroeconomic policy decisions made in each country at different points in time.

As economists, we use theory to organize and understand the data. We need to abstract from details and particular idiosyncrasies to try to unravel general patterns. It is therefore a requirement for us that the narratives for all countries follow a unified theoretical framework, which is developed in detail in section 3 of this chapter. The authors of studies of our set of eleven countries then use our theoretical framework to link the data to the sequence of the main
macroeconomic events for each of the countries and to assess the role of the different macroeconomic policies enacted, whenever possible.

We keep the conceptual framework as simple as possible, following the detailed discussion in chapter 1. We hope that the framework allows us to capture the principal forces behind the sequence of events that our narratives describe. Therefore, by construction, we are not be able to capture the effect of forces that arise only sporadically or in only a few of the countries. The theoretical framework cannot therefore be applied dogmatically. We use it systematically, but we need to acknowledge the cases in which it fails to provide a convincing explanation of the facts. The following chapters provide several examples of events that do not conform to the logic of the theory. Some qualifications to the basic conceptual framework, which we briefly discuss, can go a long way toward explaining some of the initial anomalies, as the following chapters argue. Finally, the narratives that follow also highlight events that challenge the conceptual framework and eventually will suggest avenues for further work.

The conceptual framework focuses on the relationship between the joint determination of fiscal and monetary policies and their interaction with nominal instability, as discussed above. It therefore lacks any theory of the determination of total average real economic activity. As such, the conceptual framework is unable to link economic crises with the poor economic performance that is evident in figure 1. The narratives that follow therefore limit themselves to establishing the time coincidence of economic crises with the large and persistent recessions—and even great depressions—that were so common in the region during this period. Thus, we fall short of providing a summary of policies that can lead to sustainable growth and bring prosperity to the region. Nevertheless, we hope that the database and the narratives we provide can be used by policymakers as laboratory experiments from which to draw useful lessons, from both the lost decades at the end of the twentieth century and the better decade experienced at the dawn of the twenty-first. We also hope that this sequence of studies will serve as a motivation to others in pursuing the quest for policy rules that can help to break the vicious cycle of crisis and stagnation in Latin America.

In section 2 of this chapter, we very briefly describe the macroeconomic performance of the eleven countries since 1960. We also summarize the macroeconomic instability that reigned in
the region, particularly during the interim period. Finally, in section 3, we describe the framework that serves as the guiding theoretical apparatus to organize the data and construct the narratives for each of our eleven countries.

2. Economic performance and macroeconomic instability

We now describe the evolution of the income per capita of the Latin American countries included in the study, relative to the income per capita of the world frontier. The data are from the Maddison Project Database, version 2018 (Bolt et al. 2018). We use real GDP per capita measured in 2011 US-dollars, based on multiple benchmark comparisons of prices and income across countries that make the data more suitable for cross-country income comparisons (see Bolt et al. 2018 for a more extensive explanation of the construction of this variable).

Standard practice is to use the income per capita in the United States as a proxy for the world frontier. In doing so, however, the measure is affected by idiosyncratic events in the United States, like the severity of the Great Recession and the expansion during World War II. In contrast, we use a trend growth for the United States, which grows at 2 percent per year, and which is equal to the observed income per capita in the United States in 1960. The year 1960 has the advantage that the observed value was very close to a trend computed for the period 1985–2012. That is, the 2 percent per year growth line for the United States that runs through 1960 is essentially the same line as the regression trend line for 1985–2012. We then calculate the relative income of each of the eleven countries and the regional average by dividing its observed real GDP per capita by the US trend and plot the results in figure 1.
Figure 1. Real GDP per capita relative to the United States

(a) 1935–1973
(b) 1973–2000
(c) 2000–2016

*The data for Paraguay starts in 1939, the graph assumes 2% annual growth between 1935 and 1939.

**Average, weighted by population.

Each colored line in the figure represents a different country and the solid black line represents the average for all countries, weighted by population. Panel (a) depicts the data from 1935, right after the Great Depression, to 1973, the year in which nominal instability, measured as the average volatility of the inflation rate, starts to grow (see the discussion below). The figure shows a very modest but positive convergence for the average of the region to the levels of income per capita of the United States, particularly at the end of the sample. The weighted average of the relative income per capita that was 21 percent in 1935 grew a modest 5 percent by 1973, amounting to an incremental growth of about 0.13 per year.
This convergence is the combination of three different experiences. First, there is divergence of the three initially richest countries: Argentina, Chile, and Uruguay. In addition, there is also divergence for two of the initially very poor countries: Paraguay and Bolivia. These are the only countries that had in 1973 a lower value for relative income per capita than the value they had in 1935. These effects, however, are more than compensated for by a substantial convergence of the initially middle-income countries and by Brazil, which was initially the poorest. Notice also that two remarkable success stories occur in the two largest countries measured by population, Brazil and Mexico. The third success story is Venezuela.

Panel (b) depicts the data from 1973 to 2000. Notice that the vertical axes are exactly the same as in panel (a). This figure depicts the substantial failure of Latin America to continue to develop economically. By the end of the century, the average GDP per capita had diverged back to 23 percent of the US growth trend—a number barely above the one in 1935. Had the region kept the pace of convergence of the previous period—about 0.13 per year—the ratio would have been almost 30 percent. That is equivalent to an income per capita that is about 30 percent higher than what it was in 2000—a sizable lunch. Only two countries—Brazil and Paraguay—end the period with values above the ones in 1973, and only barely so.

Finally, in panel (c), we show the data for the first years of this century. This period reflects better performance, showing a resumption of the slow convergence of Latin America to the United States to an average of 26 percent by 2016.

As mentioned above, the reasons for Latin America lagging behind in terms of economic growth still remain somewhat of a puzzle for economists. Many hypotheses have been analyzed, but no systematic and comparative analysis has been performed that could orient policy in a systematic and predictable way. A lack of understanding, however, does not mean that the problem has not been acknowledged. The revival of studies on economic growth that started in the 1980s included a series of papers, pioneered by Barro (1991), that attempted to establish empirically the existence of economic convergence, a common prediction of standard neoclassical growth models. Prominently significant in cross-country studies that included countries in Latin America was the so-called "Latin American dummy," which identified a negative effect on the growth rate between 1960 and 1985 for countries in Latin America, even when a series of controls were
added to the regressions. The systematically poor performance of the region, which is detected by the dummy in those regressions, can clearly be appreciated in figure 1 although this set of countries was significantly poorer than the United States in the early 1960s (the initial year for the Barro regressions), they had failed to maintain convergence to the United States by 1985 (the final year in those same regressions). The purpose of this book is to shed light on the underlying causes of the Latin American dummy.

Coincidental with poor growth outcomes, the lost decades of the 1970s and 1980s were the years in which the region went through the highest macroeconomic instability of its history. Latin America during that period is plagued by chronic inflation, balance of payments crises, financial crises, defaults, hyperinflations, major confiscations of assets, and bailouts of private-sector debts.

To illustrate this coincidence of poor growth outcomes with macroeconomic instability, figure 2 shows the average inflation for the eleven countries at a monthly frequency.

**Figure 2. Average monthly inflation**

![Figure showing average monthly inflation](image)

We chose to plot monthly inflation rates because, at the more standard yearly frequency, the two peaks in 1986 and 1991 blur the scale of the figure. We plot average inflation for the eleven countries included in this study, separated into two groups. The first group—whose inflation rates correspond to the scale on the left axis—includes the five countries, Argentina, Bolivia,
Brazil, Chile, and Peru, with higher average inflation; the second group—whose inflation rates correspond to the scale on right axis—includes the other six countries, Colombia, Ecuador, Mexico, Paraguay, Uruguay, and Venezuela. We also plot two vertical lines corresponding to 1973 and 2000, the years that correspond to the sub-periods in figure 1.

In figure 3 we graph rolling volatilities of the inflation rates over time. Specifically, for each country we consider windows of eight years for the inflation rate and compute, for those years, the standard deviation of the inflation rate. We then plot, for each year, the corresponding volatility, starting in 1968.

Figure 3. Rolling inflation volatility

Clearly, the worst years in terms of economic performance correspond to the periods of higher and more volatile inflation rates.

A closer look at the behavior of each country shows that five of them (Argentina, Bolivia, Brazil, Chile, and Peru) suffered one or more hyperinflations between 1973 and 1994; Venezuela is going through another one as we finish this book. Runs against the domestic currency that led to balance of payments crises were too common to count, and many countries in the region defaulted on their government debt, in several cases more than once. The wave of defaults around the world in the early 1980s that jeopardized the health of several large US banks has been named the Latin American default crisis, even though default was also declared by some
countries that are not part of Latin America. Nevertheless, it was labeled as a Latin American default crisis because most of the heavy borrowing was done by governments in that region. Massive banking crises, with bailouts that amounted to several percentage points of GDP, have occurred in almost all countries between the late 1970s and the early years of the current century. Again, some countries experienced more than one crisis. All sorts of different policies have been tried, including dual exchange rates (in every country), a currency board that ended in a major crisis (in Argentina), and full dollarization (in Ecuador), a policy that has lasted to this day. There were also several periods of floating exchange rates, a policy still prevailing today in many countries of the region. Banking crises have been dealt with in different ways in different countries and periods. Default renegotiations have also varied across countries and time periods. Banks have been nationalized and privatized.

The coincidence over time between these dramatic policy decisions, macroeconomic instability, and the economic decadence of Latin America makes its recent history a very rich experiment to analyze. As we mentioned above, however, as a profession, we still lack good theories that associate macroeconomic instability with economic performance—at least the ones that can be subject to serious quantitative scrutiny. Nevertheless, a substantial literature associates bad macroeconomic fiscal and monetary policies with macroeconomic instability: balance of payments crises, financial crises, defaults, hyperinflation, and so on. We now briefly describe a conceptual framework that summarizes the many contributions to that literature and which serve as the theoretical framework to support the eleven narratives constituting the core of this book.

3. Conceptual framework

To develop a framework to evaluate the impact of sets of monetary and fiscal policies implemented in our eleven Latin American countries over the period 1960–2017, we start by describing the framework developed by Sargent (1986) to evaluate the impact of the set of monetary and fiscal policies implemented in the United States in the early 1980s. This framework consists of two main ideas: a budget identity for the consolidated government and a demand for real money.
The budget identity classifies all sources of government financing into three groups: tax revenues, interest-bearing debt and non-interest-bearing debt, or money. As such, it imposes a constraint between four different dimensions of macroeconomic policy: total government expenses, total revenues, increases in government debt, and increases in the money supply. The constraint implies that the four different policy decisions cannot be made independently. Once three of them are decided, the fourth has to adjust to satisfy that constraint. In what follows, we combine two of the policy variables into one by netting our total revenues from total expenditures to obtain a measure of the deficit or surplus in case it is negative. We take this approach because the government budget constraint does not independently restrict spending and revenues, only its difference, as will become clear below. A direct implication of this constraint is that a deficit implies an increase in government debt, an increase in the money supply, or a combination of both.

The demand for real money establishes a systematic relationship between the general price level, short-term nominal interest rates, total real income, and some measure of money. It implies that systematic increases in the money supply generate inflation.

The combination of these two main ideas does not imply that sustained deficits cause inflation, since they can be financed by increases in government debt. Nevertheless, debt implies a promise of future government surpluses to be used to pay for that debt. To the extent that these promises lack credibility, the government may face a limit on its ability to borrow. If this is the case, the combination of the two ideas implies a direct connection between fiscal deficits and inflation. Thus, the size of interest-bearing debt relative to total production plays a key role.

As many of the cases studied illustrate, the framework just described cannot rationalize some of the crises that the region experienced during the period under study. This should not be surprising, since Sargent's object of study was the macroeconomy of the United States in the early 1980s, where moderate inflation was the only macroeconomic problem. The difference between Latin America and the United States in the early 1980s does not lie exclusively with the much higher and more volatile inflation rates observed during the period. As mentioned before, the region experienced other types of crises, and we need to adapt the framework to study them.
We first consider an open economy version of the Sargent’s (1988) framework to study balance of payments crises. We then briefly discuss models that exhibit default in equilibrium, a policy option that is not considered in the basic framework since it did not seem to be an issue for the US government, but which was a policy option chosen by several of the Latin American countries during the period under study. We then move to models in which not only the size but also the characteristics of total debt matter for the determination of equilibria. We review models where, depending on the size of short-term debt or the units in which debt is denominated, or both, can determine the outcome in the economy. In particular, we review models that address the possibility of multiple equilibria, in which a crisis may occur driven by expectations alone, in spite of the fundamentals of the model being right. The discussion of these variations is self-contained by using very simple versions of those models.

At the end of this section, we briefly mention (but do not discuss) other theoretical results derived from the theory of optimal dynamic contracts with enforcement constraints. These models complement in a natural way the basic framework and may be useful in thinking about limits to total debt. Those models raise several interesting questions that we briefly address.

3.1 The economics of budget constraints

The first building block of the conceptual framework is the government budget identity. In describing sources of financing, we separately specify, when possible, domestic currency denominated debt, inflation indexed, and foreign currency denominated debt. Specifically, we let $B_t$, $b_t$, and $B^*_t$ be total nominal, indexed, and dollar denominated debt and $D_t$ be the deficit of the governments in real terms, measured as expenditures and normal transfers minus taxes. We also let $M_t$ be the stock of money, $P_t$ be the domestic price level, that is, the GDP deflator, and $E_t$ the nominal exchange rate. Furthermore, we let $R_t$, $r_t$, and $B^*_t$ be the gross returns on nominal, inflation indexed, and foreign currency bonds. Then the budget constraint of the government is

$$B_t + P_t b_t + E_t B^*_t + M_t = P_t (D_t + T_t) + B_{t-1} R_{t-1} + P_{t-1} b_{t-1} r_{t-1} + E_{t-1} B^*_{t-1} R^*_{t-1} + M_{t-1}$$

(1)
Unless explicitly mentioned, the stock of debt, $B_t$, does not include the assets and liabilities of the central bank. These can be important for some countries and in some periods of time, and will be mentioned explicitly in the case studies that follow.

Notice that on the right-hand side of equation (1), we have added a term $T_t$ to the deficit. We do this because we have independent measures for all the other terms in the equation, and we can measure $T_t$ as a residual. Specifically, we choose the value for $T_t$ that makes the budget constraint hold, given the values for all the other terms. In many cases, the variable $T_t$ allows us to identify off-the-book expenses, particularly in times of crisis. To fix ideas, we can provide two examples drawn from the experience of several countries. Following banking crises, governments in many circumstances chose to bail out the financial sector, typically by issuing government bonds. These would be accounted for as a positive value for $T_t$. In some cases, the government also provided subsidies through state-owned development banks or state-owned companies and, in many occasions, used seigniorage from the central bank to cover those losses. These off-the-book expenses would also show up as positive values for $T_t$. As many of the chapters in this book illustrate, the accumulated effect of these computed transfers can be very large over time. In some cases, the narratives help to identify the economic forces that created these large transfers.

The real exchange rate is

$$
\xi_t = E_t P^W_t \frac{P_t}{y_t}
$$

where $P^W_t$ is the price level for dollar denominated debt, that is, the US GDP deflator. Then,

$$
E_t \left( \frac{B_t^*}{P^W_t} \right) \frac{P_t}{y_t} = E_t P^W_t \left( \frac{B_t^*}{P_t} \right) \frac{P_t}{y_t} = E_t \frac{B_t^*}{P_t y_t}
$$

is the value of dollar denominated debt as a fraction of nominal GDP, $P_t y_t$.

If we let
\[
\theta_t^n = \frac{B_t}{P_t y_t}, \quad \theta_t^r = \frac{b_t}{y_t}, \quad \theta_t^* = \frac{\theta_t^r / P_t^*}{y_t}, \quad m_t = \frac{M_t}{P_t y_t}, \quad d_t = \frac{P D_t}{P_t y_t}, \quad \tau_t = \frac{P T_t}{P_t y_t}
\]

and

\[
\xi_t = \left( \frac{E_t P_t^*}{P_t} \right), \quad g_t = \frac{Y_t}{Y_{t-1}}, \quad \pi_t = \frac{P_t}{P_{t-1}}, \quad \pi_t^w = \frac{P_t^w}{P_{t-1}^w}.
\]

we can write the budget constraint in terms changes as fractions of GDP as

\[
\left( \theta_t^n - \theta_{t-1}^n \right) + \left( \theta_t^r - \theta_{t-1}^r \right) + \left( \xi_t \theta_t^* - \xi_{t-1} \theta_{t-1}^* \right) + \left( m_t - m_{t-1} \right) + m_{t-1} \left( 1 - \frac{1}{g_t \pi_t} \right)
\]

\[
= d_t + \tau_t + \theta_{t-1}^n \left( \frac{R_{t-1}}{g_t \pi_t} - 1 \right) + \theta_{t-1}^r \left( \frac{r_{t-1}}{g_t} - 1 \right) + \xi_{t-1} \theta_{t-1}^* \left( \frac{R_{t-1}^w}{g_t} - 1 \right).
\]

(A detailed derivation of this equation is in Appendix A.) The first three terms on the left-hand side of our budget accounting equation (4) represent increases in debt-to-output ratios in the three different types of debt: nominal, indexed, and foreign currency. The fourth term represents increases in high-powered money, and the last one measures seigniorage. The first three terms on the right-hand side represent the service costs on each of the three debt types. The final two terms represent the fiscal deficit—including the extraordinary transfers \( T_t \) — as a fraction of output.

All eleven of the countries that we study in this volume imposed dual or multiple exchange rates during the 1970s or 1980s or both. What exchange rate should we use for the nominal exchange rate \( E_t \) in the budget constraint (1) and to construct the real exchange rate \( \xi_t \) in equation (2)? Ideally, we should use the rate that corresponds most closely to a market rate. If there is a real devaluation, using the market rate—rather than a lower official rate or preferential rate—better captures the magnitude and the timing of the increase in the burden of foreign debt implied by the devaluation. To the extent that the central bank exchanges foreign currency for local currency at the lower official rate for some importers, for example, it is subsidizing purchases of imports by these agents. Conversely, to the extent that the central bank exchanges local currency for foreign currency for some exporters at the official rate, it is taxing the exports by these agents. Typically, we do not have information on the purchases or sales of foreign currency at different
exchange rates, which means that these implicit subsides and taxes are included in the transfer term.

The second building block is a demand for real money balances, which we write as

$$\frac{M_t}{P_t} = \gamma - \delta \left( \frac{P_{t+1}}{P_t} \right)$$

where $M_t$ is the outstanding stock of money, $P_t$ is the price level, and $\gamma > \delta$. This money demand equation arises in a simple overlapping-generations general equilibrium model, as the models described in Sargent (1986) and Marcet and Nicolini (2003) and as we show in Appendix B. In a stochastic model, we would put replace the inflation factor $P_{t+1}/P_t$ with its expected value, but here we keep our discussion simple by focusing on the deterministic model. The linearity of the money demand equation is not essential to our arguments, but makes the discussion simple. Notice that neither real output nor the real interest rate appears in equation (5), because in the above-mentioned models, both of those variables are constant over time and are embedded in the parameters $\gamma$ and $\delta$. The assumption of constant real output is without loss of generality, since the variable $M_t$ can be interpreted as the ratio of money balances to output. The assumption of a constant real interest rate is less innocuous, but nothing essential in what follows hinges on that assumption.

We can solve forward the difference equation defined in (5) and write the unique non bubble solution for the price level as

$$P_t = \frac{1}{\gamma} \sum_{j=0}^{\infty} \left( \frac{\delta}{\gamma} \right)^j M_{t+j}.$$ 

This equation implies that sustained increases in money growth generate sustained increases in prices.

The government budget constraint implies that sustained deficits lead either to sustained increases in the quantity of money or to sustained increases in government debt. The first option implies sustained inflation, as it does for the money demand equation. The second option implies that the government may eventually face constraints in its ability to borrow.
Fiscal deficits do not necessarily imply that inflation needs to increase because of the possibility of issuing bonds. As does Sargent (1986), however, we assume that there is a limit on the ability to borrow, which, as a first approximation, should be related to the ability of the government to generate future surpluses. Specifically, we assume that there is a debt limit,

\[
\theta_i^n + \theta_i^f + \xi_i \theta_i^f \leq \Theta, \tag{7}
\]

where \( \Theta \) is an exogenously specified number. We summarize some models that attempt to understand the value of \( \Theta \) below.

To the extent that the debt constraint (7) is not binding, governments can use fiscal policy as a tool to smooth shocks that affect the business cycle, without creating inflation. All adjustments required to satisfy the budget constraint can be done by properly managing the debt. This regime is a very good approximation to the way the governments in the United States and the United Kingdom financed the war efforts in the last century. (See, for example, Hall and Sargent 2011.) These circumstances may appear to be ones in which monetary and fiscal policy can be designed independently from each other. This is not true, however: debt financing does not eliminate the interdependence between fiscal and monetary policy; it only postpones the debate. This is the central message of the analysis in Sargent (1986). While debt financing is available, deficits may not interfere with the design of monetary policy. Eventually, however, once the debt is high enough—that is, the debt constraint (7) is binding—the debate naturally arises and a game of chicken arises between the monetary authority and the fiscal authority.

This situation is characterized by fiscal dominance when the winner of the game of chicken is the fiscal authority and the central bank monetizes the deficits. The reverse case, in which the fiscal authority generates surpluses, is one of monetary dominance. As it turns out, in the United States, monetary dominance prevailed as fiscal surpluses in the 1990s followed the Reagan deficits. In contrast, the dramatically high inflation rates observed in Latin America are evidence of all-too-frequent periods of fiscal dominance. High inflation rates followed once a specific country had already reached its debt limit and the financing requirements, given by the right-hand side of equation (4), were positive. In these cases, the budget constraint becomes
\[
\left( m_t - m_{t-1} \right) + m_{t-1} \left( 1 - \frac{1}{g_t \pi_t} \right) \\
= d_t + \tau_t + \theta_{t-1}^M \left( \frac{R_{t-1}}{g_t \pi_t} - 1 \right) + \theta_{t-1}^N \left( \frac{r_{t-1}}{g_t} - 1 \right) + \xi_t \theta_{t-1}^* \left( \frac{R_{t-1}}{g_t \pi_t^*} - 1 \right) \tag{8}
\]

and inflation is unavoidable.

The debt constraint (7) would naturally bind during the years immediately following a default, as it did for many countries in Latin America at the beginning of the 1980s. A direct implication of this analysis is that the countries that defaulted and ran deficits in the following years ought to experience inflation. This is indeed the case in the studies that follow.

This positive relationship between deficits and inflation rates over time is a direct implication of the model above, which can be solved using perfect foresight: notice that the future price level in the money demand equation is the same as in the equilibrium price level. As shown in Marcet and Nicolini (2003) and Sargent, Williams, and Zha (2009), however, the joint equilibrium dynamics of the deficit and the inflation rate can be different if one allows for small departures from rational expectations. Interestingly, the model dynamics in these papers closely resemble many features of the data. We ignore those details in the narratives that follow. The interested reader can consult the papers just mentioned for details.

As we have mentioned, Latin American experiences feature events that are apparently much more complicated than those of the United States experience, for which Sargent (1986) designed the conceptual framework. We now describe some relatively minor modifications to the framework that will help us interpret those events.

### 3.2 Balance of payments crises

The framework above can also accommodate balance of payments crises when countries chose to fix the devaluation rate, as most countries in Latin America did at some points during the period that we study. We briefly describe how to introduce the sort of balance of payments crisis studied by Krugman (1979).

We assume that purchasing power parity holds:
\[ P_t = E_t P_t^w \]  

By successfully fixing the rate of depreciation of the nominal exchange rate, \( E_{t+1} / E_t \), and given foreign inflation, the government pins down the domestic inflation rate. Under these conditions, the money demand equation (5) determines the path for the money supply.

Consider now a country whose debt has reached its debt constraint (7) and which chooses to fix the devaluation rate. The relevant budget constraint in this case is given by equation (8), where the left-hand side is given by the evolution of inflation together with the money demand equation, as explained above. Given a value for the primary deficit and the interest payments, the only variable left to satisfy the government budget constraint (8) is the transfer \( t_t \). The natural interpretation is that the reserves at the central bank adjust to satisfy the budget constraint, and this is one rationale for the key role played by the central bank's stock of reserves in fixed exchange rate regimes.

A balance of payments crisis unfolds after a sequence of positive deficits that are financed with those reserves. While this occurs, the exchange rate regime suppresses domestic inflation through the mechanism described above at the cost of a systematic decline in the stock of reserves. Eventually, agents foresee that, if there were a speculative attack, the central bank would not have enough reserves to support the exchange rate regime. A devaluation ensues, which, through the budget constraint (8), pushes domestic prices upward. Thus, exchange rate regimes may delay the inflationary consequences of chronic deficits. The delay is paid for by the reduction in the stock of foreign reserves held at the central bank. Again, the inability to borrow is a key component of the theory.

3.3 Equilibrium default

The model above has an exogenous borrowing constraint, so to the extent to which that constraint binds for a particular country, access to the debt market is restricted, and the connection between deficits and inflation becomes tight. In that model, however, default never occurs.
A large literature developed following the contributions of Aguiar and Gopinath (2006) and Arellano (2008), who build on Eaton and Gersovitz (1981), to address that issue. This literature assumes that government debt is non-contingent and that governments are sovereign in the sense that they cannot commit to repaying so default is always an option, and there is either no or very limited ability to collateralize the debt. Default is assumed to be costly in terms of lost output so the models do exhibit default in equilibrium. The implications of these models will be taken into account in the narratives that follow, and the data and the narratives can narrow down the theories and discipline the parameters to deepen our understanding of these dramatic episodes.

3.4 The maturity problem

More recent papers, developed after the 1994–1995 Mexican crisis, such as Calvo (1998) and Cole and Kehoe (1996, 2000), have emphasized the maturity structure of debt rather than the total value of the debt. These papers develop models in which debt crises can be touched off by the expectations of investors in government bonds—that this, investors’ expectations of a crisis are self-fulfilling—but in which the possibility of such crises can be reduced or eliminated by having debt of a long enough maturity.

To understand how these models work, we consider a simple two-period economy based on Calvo (1998). The government inherits debt $B_1$ and has the budget constraint

$$ B = s_1 + \frac{s_2}{R} $$

where $s_t = -d_t$ is the primary surplus of period $t$, and $R$ is the (gross) international interest rate. The debt is positive, so the government will need positive surpluses to pay it back.

In this simple model, we impose three assumptions to ensure expectations-driven multiplicity. First, we assume away any enforcement problems. Second, we assume that $\bar{s}$ is the maximum surplus that the government can raise without provoking a recession. Finally, we assume that a recession means that in the following period, no surplus can be raised, and consequently the government must default on any remaining debt.
We now develop conditions on $B$, $R$, and $s$ for which there is no debt crisis if investors do not expect a crisis but there is a crisis if investors expect one. Given our assumptions, if

$$B \leq \bar{s},$$

then all the debt can be paid in the period 1, and a recession can be avoided in both periods. That is one way to pay for the debt, but clearly, any pair of surpluses that satisfies (10) can also achieve that goal without creating a recession.

On the other hand, if

$$B > \bar{s} + \frac{\bar{s}}{R}$$

then the debt cannot be paid without avoiding a recession. A debt crisis necessarily occurs in the first period and the government defaults.

Now we consider the more interesting case of intermediate values of the debt:

$$\bar{s} < B \leq \bar{s} + \frac{\bar{s}}{R}$$

(11)

Imagine that all the debt is due in the first period. This case has two possible equilibria. In the first equilibrium agents there is a positive surplus in the first period, close enough to, but less than or equal to $\bar{s}$, that covers part of the debt, while a share of the debt is refinanced for repayment in period 2. Then, a surplus can be generated in period 2 that is enough to pay back all the remaining debt. Clearly, both of the surpluses can be less than or equal to $\bar{s}$, so recessions are avoided. This equilibrium depends on investors being willing to refinance $B - s_1$ from period 1 to period 2.

In the second equilibrium, none of the debt gets refinanced, so the government is forced to raise a surplus that is higher than $\bar{s}$. This provokes a recession implying that the government in the second period is unable to raise a surplus, making it rational for the lenders to not refinance the debt. In this second equilibrium, the government may be forced to default in the first period if it is unable to raise a surplus that is large enough to pay all the debt. Whether, the government defaults in period 1 and suffers whatever default penalty we specify or it generates the surplus
necessary to pay back the debt and suffers the recession depends on how we specify the costs of recessions and defaults.

This case is interesting because the maturity of the debt can be managed to eliminate the second equilibrium. If an amount equal to $\bar{s}$ is due in the second period, then the government has automatic refinancing, does not need to raise a surplus larger than the maximum, and thus avoids a recession.

The model that we have outlined has two equilibria. We can think of the determination of which equilibrium occurs as being determined by a randomization device, a sunspot. Cole and Kehoe (1996, 2000) develop an infinite horizon model with sunspots. This model has richer implications but is more complex. The probability of a negative sunspot now determines a risk premium on government debt, and this risk premium feeds back into the determination of the crisis zone—the set of debt levels that satisfy the analogue of condition (11)—that is, the set of debt levels for which there are a repayment continuation equilibrium and a default continuation equilibrium. The higher the risk premium, the more difficult the government finds it to pay back its debt even if there is no negative sunspot, and the crisis zone shrinks. Cole and Kehoe (1996) also show that the longer the maturity of debt, the smaller the crisis zone. In the case of debt in perpetuities, the crisis zone disappears. The lesson from analyzing high probability crises and long maturity debt is the same: It is not the total amount of debt that is crucial in determining whether or not a crisis can occur, but the amount of debt service required every period.

3.5 The denomination problem

The chapters that follow are replete with stories of Latin American governments having debt crises because they issue debt denominated in dollars. We can use a simple version of alternative multiple equilibrium developed by Calvo (1988), however, to illustrate the benefits of issuing debt denominated in a foreign currency like US dollars. Assume that a government launches a stabilization plan that fixes the nominal exchange rate. For simplicity, assume that the foreign interest rate is fixed at $R^*$. Then, if the investors who buy the bonds expect the government to keep the exchange rate fixed, then the expected devaluation is zero, and the domestic interest rate is also $R^*$. 

21
This government can modify fiscal policy to generate primary fiscal surpluses, but some uncertainty is involved. In particular, we assume the surplus will be a random variable, and, to keep our discussion simple, we assume that it can take on two values, \( s_t = s < 0 \) with probability \( \pi \) and \( s_t = \bar{s} > 0 \) with probability \( 1 - \pi \).

Given the stock of government debt, \( B_t \), the government’s financing needs are

\[
B_t R_t - s_t \in \{B_t R_t - \bar{s}, B_t R_t - s\}.
\]

We again assume that there is a maximum \( \bar{B} \) that the market is willing to lend to this government in a given period. If financing needs in a given period are larger than this maximum, then we assume that the government has to use its international reserves, a balance of payments crisis occurs, and the government then devalues at the rate \( e = E_t / E_{t-1} > 0 \).

If the maximum \( \bar{B} \) is larger than \( B_t R^* - \bar{s} \), then there is an equilibrium with \( R_t = R^* \) and no devaluation. In fact, if investors expect devaluation to be zero, then the probability that financing needs will be higher than the maximum is zero, and thus the probability of a devaluation is zero. On the other hand, if investors expect a balance of payments crisis to occur with positive probability, then they demand a higher interest rate on the bonds, \( R_t > R^* \). In the equilibria that we examine, investors either assign devaluation the probability 0 or the probability \( \pi \) because they expect devaluation to occur only if the government runs a deficit. It is worth mentioning that, if the maximum \( \bar{B} \) is smaller than \( B_t R^* - \bar{s} \), then the unique equilibrium is one in which the government devalues.

If investors are risk neutral, they require that the expect discounted value in period \( t \) of the bonds that they bought in period \( t-1 \) to satisfy the arbitrage equation

\[
1 = \beta \left( \frac{R_t}{e} + (1 - \pi)R_t \right)
\] (12)

Notice that, if \( \pi = 0 \), then the country can borrow at the risk-free rate \( R_t = R^* = 1 / \beta \) as we have explained. If, however, investors expect the government to devalue with probability \( \pi \), this the arbitrage equation (12) tells us that
Can risk premium $R_t - R^*$ in equation (13) be part of a rational expectations, self-fulfilling balance-of-payments crisis equilibrium? For this to be the case, we require that

$$B_t R_t - s > B_t R^* - s$$

(14)

How large is the crisis zone defined by condition (14) depends on the parameters $\bar{B}_t, s, \mu$, and $e$. In this case, the multiplicity is due to a denomination problem: since the debt is denominated in domestic currency and a devaluation can reduce the value of the debt, there is a low interest rate equilibrium, where the debt maintains its value with probability one, and a high interest rate equilibrium, where the debt devalues with positive probability.

More recently, Lorenzoni and Werning (2013) show that a similar type of multiplicity can arise with debt of long maturities. Ayres et al. (2018) do the same in models in which the economy faces the likelihood of relatively long periods of stagnation.

These models suggest that the total size of government debt may not be the unique determinant of the ability of the government to borrow and that market expectations can play an independent role. The extent to which these theoretical considerations can explain some of the default episodes during the period will be addressed in the studies that follow.

3.6 Enforcement problems

In this subsection, we mention theoretical developments that address the issue of willingness to pay. This feature is crucial, since most government debt does is not collateralized. In the absence of commitment, the government must find it optimal to repay the debt every period. A partial list of papers in this area include Albuquerque and Hopenhayn (2004) and Phelan (2006). In some of the equilibria in these models, the maximum amount that the government is able to borrow depends on the history of shocks and changes over time. Many times, the borrowing constraint is either binding or close to the bound. It is therefore possible that, even if a government has a
debt-to-GDP ratio that is relatively low, it could be unable to borrow a few percentage points of GDP in a single period.

There is a sense in which some of these models could rationalize the distinction between "emerging" governments—the ones that only recently entered the international capital markets—and "developed" governments—the ones that have used the debt markets for a long time. In these setups, it is interesting to reconsider the liberalization of capital markets and financial systems. Experience shows that liberalizing the financial sector substantially increases contingent debt, and if a shock is realized that causes the government to bail out banks, then the need for debt runs discontinuously high, which may not be consistent with the credit constraints of the problem. The discussion suggests that a good part of the debt crisis in the early 1980s may be explained by the crackdown in financial sectors and the substantial increase in government debt, due to the deposit insurance. Thus, the combination of an emerging government, together with a liberalization of capital flows may be an explosive combination because the emerging government could find itself limited in the amount of credit that it can obtain in the market. The emerging nature of the government implies that it will be credit constrained for a time. On the other hand, financial liberalization substantially increases the contingent debt. Tension seems to be present between opening the country to foreign capital and early financial liberalization. This is the central message of Diaz-Alejandro (1985).

These considerations will be relevant in trying to understand some of the crises experienced by the region during this period. They may help to shed light on questions such as, Why there is a debt limit to begin with? Does it make a difference if you arrive at that limit smoothly or by a discrete jump in the debt? To put it differently, should the constraints on debt be related to the total amount of debt in a given period or to the net change in the total amount of debt in a given period? How relevant are incentives rather than ability to repay in determining credit limits? We will fall short of providing clear answers to these difficult and very important questions, but we are convinced that the case studies that follow bring us closer to these answers.
References


Appendix A

The government’s budget constraint in units of domestic currency is

\[ B_t + P_t b_t + E_t B_t^* + M_t = P_t (D_t + T_t) + B_{t-1} R_{t-1} + P_t b_{t-1} r_{t-1} + E_t B_{t-1}^* R_{t-1} + M_{t-1}, \]

where the notation has been defined in Section 3.1.

Dividing the equation by GDP in current prices, we obtain

\[ \frac{B_t}{P_t y_t} + \frac{P_t b_t}{P_t y_t} + \frac{B_t^* E_t}{P_t y_t} + \frac{M_t}{P_t y_t} = \frac{P_t D_t}{P_t y_t} + \frac{P_t T_t}{P_t y_t} + \frac{B_{t-1} R_{t-1}}{P_t y_t} + \frac{P_t b_{t-1} r_{t-1}}{P_t y_t} + \frac{E_t B_{t-1}^* R_{t-1}}{P_t y_t} + \frac{M_{t-1}}{P_t y_t}. \]

We can rewrite this equation as

\[ \frac{B_t}{P_t y_t} + \frac{P_t b_t}{P_t y_t} = \frac{(E_t P_t^W)^{y_t}}{P_t y_t} B_t^* / P_t^w + \frac{M_t}{P_t y_t} \]

\[ = \frac{P_t D_t}{P_t y_t} + \frac{P_t T_t}{P_t y_t} + \left( \frac{P_t^w y_t - 1}{P_t y_t} \right) \frac{B_{t-1} R_{t-1}}{P_t y_t} + \left( \frac{P_t^w y_t - 1}{P_t y_t} \right) \frac{P_t b_{t-1} r_{t-1}}{P_t y_t} + \left( \frac{E_t P_t^w}{P_t^w y_t} \right) \frac{B_{t-1}^* R_{t-1}}{P_t y_t} + \frac{M_{t-1}}{P_t y_t}. \]

Now let

\[ \theta_t^N = \frac{B_t}{P_t y_t}, \quad \theta_t^F = \frac{b_t}{y_t}, \quad \theta_t^* = \frac{B_t^* / P_t^w}{y_t}, \quad \theta_t = \frac{M_t}{P_t y_t}, \quad d_t = \frac{P_t D_t}{P_t y_t}, \quad \tau_t = \frac{P_t T_t}{P_t y_t}, \]

and

\[ \xi_t = \left( \frac{E_t P_t^w}{P_t^w} \right), \quad g_t = \frac{y_t}{y_{t-1}}, \quad \pi_t = \frac{P_t}{P_{t-1}}, \quad \pi_t^w = \frac{P_t^w}{P_{t-1}^w}. \]

We can write the budget constraint as

\[ \theta_t^N + \theta_t^F + \xi_t \theta_t^* + \theta_t = d_t + x_t + \frac{R_t-1}{g_t \pi_t} \theta_t^N + \frac{r_t-1}{g_t} \theta_t^N + \frac{R_t^*}{g_t \pi_t^w} \theta_t^* + \frac{1}{g_t \pi_t} m_{t-1}. \]

Subtracting \( \theta_t^N, \theta_t^F, \xi_t \theta_t^*, \) and \( m_{t-1} / (g_t \pi_t) \) from both sides of this equation, we obtain

\[ \left( \theta_t^N - \theta_{t-1}^N \right) + \left( \theta_t^F - \theta_{t-1}^F \right) + \left( \xi_t \theta_t^* - \xi_t \theta_{t-1}^* \right) + \left( \theta_t - \theta_{t-1} \right) + m_{t-1} - \frac{1}{g_t \pi_t} m_{t-1} \]

\[ = d_t + \tau_t + \theta_t^N \left( \frac{R_t-1}{g_t \pi_t} - 1 \right) + \theta_t^N \left( \frac{r_t-1}{g_t} - 1 \right) + \xi_t \theta_t^* \left( \frac{R_t^*}{g_t \pi_t^w} - 1 \right). \]
which is equivalent to our budget accounting equation (4) since

\[
m_t - \frac{1}{g_t \pi_t} m_{t-1} = (m_t - m_{t-1}) + m_{t-1} \left(1 - \frac{1}{g_t \pi_t}\right)
\]
Appendix B

In this appendix, we solve a deterministic small open economy version of a simple, two-period-lived overlapping-generations model in which there is a demand for money like that is equation (5).

Households

Each cohort has a unit mass of household who live for two periods. There is one type of consumption good in every period. The utility function of the representative household of the generation born in period $t$ is

$$\log c_t' + \lambda \log c_{t+1}'$$  \hspace{1cm} (15)

where $c_t'$ is the consumption of this household when young, and $c_{t+1}'$ is the consumption when old. The representative household is endowed with 1 unit of the good when young and $e$ consumption when old, where $1 > e > 0$.

Asset Markets

There are two assets in the economy: domestic and foreign currency. To ensure that there are equilibria in which domestic currency is used, we impose a cash-in-advance constraint for local currency on net purchases of consumption:

$$M_t \geq P_{t+1} (c_{t+1}' - e)$$  \hspace{1cm} (16)

This condition makes foreign currency value less for households than foreign currency. To finance expenditures, the government levies lump sum taxes $\tau_t$ on the young in generation $t$.

The budget constraint of the household born in $t$ when young is

$$P_t c_t' + M_t = P_t - P_t \tau_t$$  \hspace{1cm} (17)

The household maximizes utility (15) subject to budget constraint (17), the cash-in-advance constraint (16) and a non-negativity constraint on money holdings $M_t \geq 0$. The solution is
\[
\begin{cases}
\frac{p_t (1 - \tau_t) + p_{t+1} e}{(1 + \lambda) p_t} & \text{if } \frac{p_{t+1}}{p_t} \leq \frac{\lambda(1 - \tau_t)}{e} \\
1 & \text{if } \frac{p_{t+1}}{p_t} > \frac{\lambda(1 - \tau_t)}{e}
\end{cases}
\]

\[
\begin{cases}
\frac{\lambda(p_t (1 - \tau_t) + p_{t+1} e)}{(1 + \lambda) p_{t+1}} & \text{if } \frac{p_{t+1}}{p_t} \leq \frac{\lambda(1 - \tau_t)}{e} \\
e & \text{if } \frac{p_{t+1}}{p_t} > \frac{\lambda(1 - \tau_t)}{e}
\end{cases}
\]

\[
\frac{M_t}{p_t} = \begin{cases}
\frac{\lambda(1 - \tau_t)}{1 + \lambda} - e \left( \frac{p_{t+1}}{p_t} \right) & \text{if } \frac{p_{t+1}}{p_t} \leq \frac{\lambda(1 - \tau_t)}{e} \\
0 & \text{if } \frac{p_{t+1}}{p_t} > \frac{\lambda(1 - \tau_t)}{e}
\end{cases}
\]

Notice that the money demand function (5) has exactly this form where

\[
\gamma = \frac{\lambda(1 - \tau_t)}{1 + \lambda}, \quad \delta = \frac{e}{1 + \lambda},
\]