ESSAYS IN INTERNATIONAL FINANCE
No. 188, November 1992

PARALLEL CURRENCY MARKETS
IN DEVELOPING COUNTRIES: THEORY, EVIDENCE, AND POLICY IMPLICATIONS

PIERRE-RICHARD AGÉNOR

INTERNATIONAL FINANCE SECTION
DEPARTMENT OF ECONOMICS
PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY
ESSAYS IN INTERNATIONAL FINANCE

ESSAYS IN INTERNATIONAL FINANCE are published by the International Finance Section of the Department of Economics of Princeton University. The Section sponsors this series of publications, but the opinions expressed are those of the authors. The Section welcomes the submission of manuscripts for publication in this and its other series. Please see the Notice to Contributors at the back of this Essay.

The author of this Essay, Pierre-Richard Agénor, is an economist in the Research Department of the International Monetary Fund. His recent work has focused on the macroeconomic implications of informal financial markets in developing countries, the theory of balance-of-payments crises, and the role of credibility factors in the effectiveness of stabilization programs.

PETER B. KENEN, Director
International Finance Section
ESSAYS IN INTERNATIONAL FINANCE
No. 188, November 1992

PARALLEL CURRENCY MARKETS
IN DEVELOPING COUNTRIES: THEORY, EVIDENCE, AND POLICY IMPLICATIONS

PIERRE-RICHARD AGÉNOR

INTERNATIONAL FINANCE SECTION
DEPARTMENT OF ECONOMICS
PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY
# CONTENTS

1 INTRODUCTION

2 THE SCOPE AND NATURE OF PARALLEL CURRENCY MARKETS
   Emergence of Parallel Markets 2
   Supply and Demand for Foreign Currency 7
   Welfare Implications of Parallel Markets 10

3 DETERMINANTS OF PARALLEL EXCHANGE RATES: THEORY 13

4 POLICY ISSUES WITH PARALLEL CURRENCY MARKETS
   Effectiveness of Exchange Controls 17
   Devaluations and Parallel-Market Premia 18
   Unification of Foreign-Exchange Markets 23

5 CONCLUDING REMARKS 30

REFERENCES 31
FIGURES

1 Parallel-Market Premia in Developing Countries 4
2 Parallel-Market Premia in Unification Attempts 28

TABLE

1 Changes in Exchange Rates, Money Supply, and Prices in Nine African Countries 27
PARALLEL CURRENCY MARKETS
IN DEVELOPING COUNTRIES: THEORY, EVIDENCE,
AND POLICY IMPLICATIONS

1 Introduction

There has been growing recognition over the past few years that widespread exchange and trade restrictions in developing countries have been ineffective in preserving reserves or in supporting an overvalued exchange rate. Evasion has been endemic, and illegal markets for goods and foreign currencies have expanded, defeating the very purpose of controls. Although the nature of parallel markets precludes collection of detailed and reliable data, these markets appear to be common phenomena in developing countries, with parallel exchange rates deviating from official rates considerably in some cases.¹

This essay reviews recent theoretical and empirical analyses of parallel currency markets in developing countries and examines key policy issues related to these markets. Section 2 examines the scope and nature of these markets and highlights the basic structural characteristics likely to be found in a variety of institutional settings. Section 3 discusses the determinants of parallel exchange rates emphasized by the recent theoretical literature. Section 4 considers some policy issues faced by countries with a sizable parallel currency market. The analysis focuses, in particular, on the rationale and effectiveness of exchange restrictions, on the role of nominal devaluations as an instrument to reduce the spread between the official and parallel rates, and on strategies for unifying official and parallel markets. Finally, Section 5 provides concluding remarks and highlights some of the directions in which the existing theoretical and empirical literature can be extended.

¹ According to data presented in the World Currency Yearbook, 1989, parallel currency markets exist in all developing countries, except the high-income oil exporters. The evidence available suggests that parallel markets have recently increased in size and sophistication in many countries in relation to capital movements.
2 The Scope and Nature of Parallel Currency Markets

Due to the often illegal—albeit perhaps officially tolerated—nature of transactions in parallel markets, information on their functioning is neither readily available nor very reliable. Magnitudes mentioned here should therefore be treated with caution. The major qualitative features of parallel markets are well documented, however, suggesting that common features are to be found in a variety of institutional settings. This section discusses the ways in which parallel markets emerge, the nature of transactions conducted in those markets, and their welfare implications.

Emergence of Parallel Markets

Parallel markets generally develop in conditions of excess demand for a commodity subject to legal restrictions on sale, official price ceilings, or both. In a large majority of developing countries, transactions in foreign exchange are subject to both kinds of restrictions (see the IMF’s Annual Report on Exchange Restrictions). Typically, the exchange rate is officially pegged by the central bank, and only a small group of intermediaries is permitted to engage in currency transactions. Purchases of foreign currencies by domestic agents are, in principle, restricted to uses judged by the authorities to be “essential” for economic development, such as imports of capital goods. As a consequence, some of the supply of foreign exchange is diverted and sold illegally, at a market price higher than the official price, to satisfy the excess demand. The amount by which the parallel-market exchange rate exceeds the official rate, the “parallel-market premium,” will depend upon a host of factors—in particular, the penalty structure and the volume of resources devoted to apprehension and prosecution of violators.

Figure 1 shows the evolution of the parallel-market premium in a group of developing countries during the nineteen eighties. It shows

---

2 The expressions “parallel,” “fragmented,” “informal,” “black” (which has an illicit connotation), and “curb” markets have been used interchangeably in the literature. Lindauer (1989) provides an analytical distinction between these descriptions of market structure. He defines a parallel market (p. 1873) as “the structure generated in response to government interventions that create a situation of excess supply or demand in a particular product or factor market.” As government price fixing (through taxes, regulations, and prohibitions) plays a prominent role in the creation of excess demand at official prices, it is thus important for the emergence of a parallel market. See also Feige (1989).

3 Parallel exchange rates are taken from the World Currency Yearbook (formerly Pick’s Currency Yearbook); official exchange rates are from the International Monetary
that the premium typically displays large fluctuations over time and across countries—a phenomenon often seen as reflecting the asset-price characteristics of the parallel rate. In periods characterized by uncertainty about macroeconomic policies or unstable political and social conditions, parallel-market rates tend to react swiftly to expected future changes in economic circumstances. Figure 1 also shows that the premium has at times been substantially negative in some countries—a somewhat surprising fact because exchange restrictions in the official market relate typically to purchases of foreign currency and not to sales. Although it is difficult systematically to rationalize episodes of negative premia, a number of cases can be accounted for by three factors. First, in outward-oriented economies experiencing high rates of growth and large external surpluses (notably those in Asia), the central bank has at times restricted the rate of accumulation of foreign exchange by the banking system; capital controls have been used to inhibit the inward flow of funds, leading to periods of temporary excess supply in the parallel market. Second, periods during which a significantly negative premium has emerged may have been associated in some countries with expectations of a revaluation of the official exchange rate. Third, a negative premium may have emerged during periods when commercial banks have been forbidden to buy foreign currency without proper identification of the seller; in such circumstances, a negative premium represents a “laundering charge” (Dornbusch et al., 1983) paid by agents who have no legal right to the foreign currency they are offering for sale.

Parallel markets in developing countries typically emerge out of restrictions on foreign trade and capital flows. In low-income countries, the process often starts with the government trying to impose regulations on trade flows (licensing procedures, administrative allocations of

---

Fund database: Data are end-of-period rates relative to the U.S. dollar.

4 Figure 1 also suggests the existence of a clear seasonal pattern for some countries (Malaysia and Morocco, for instance). At a more formal level, Akgiray, Booth, and Seifert (1988) provide a statistical analysis of the distributions of parallel-market exchange rates for twelve Latin American currencies. See also Akgiray, Aydogan, and Booth (1990).

5 Small, negative premia can often be attributed to measurement errors. Data problems, however, cannot explain a large and sustained negative exchange-rate differential.

6 Trade restrictions seem to be the main factor in low-income countries. Capital controls—often motivated by recurrent balance-of-payments difficulties—are the primary factors in middle-income countries (particularly in Latin America). See, for example, Kamin’s (1991a) account of Argentina in the 1930s.
FIGURE 1
PARALLEL-MARKET PREMIA IN DEVELOPING COUNTRIES
(in percentages)
FIGURE 1 (Continued)

Morocco

Nigeria

Pakistan
foreign exchange, prohibitions). The imposition of tariffs and quotas creates incentives to smuggle and fake invoices by creating an excess demand for imported goods at illegal, pre-tax prices (Bhagwati, 1978, pp. 64-81, and see Cooper, 1974, on Indonesia). Illegal trade creates a demand for illegal currency, which, in turn, stimulates its supply and leads to the creation and establishment of a parallel currency market if the central bank is unable, or unwilling, to meet all the demand for foreign exchange at the official exchange rate. At a later stage, the parallel market expands to accommodate financing of capital flight and portfolio transactions; foreign-currency holdings are used to hedge against adverse political change and, in high-inflation economies, to hedge against the inflation tax. Many other factors may help explain the development of a parallel currency market in a particular country. In Pakistan, for instance, the rapid expansion of the illegal market for foreign exchange in the late 1970s is ascribed primarily to the sudden influx of worker remittances from the Middle East (Banuri, 1989). In Colombia and Guyana, the expansion of the illegal market for U.S. dollars has been closely associated with drug-related activities (Thomas, 1989).

Whatever the initial factors leading to the emergence or expansion of a parallel market for foreign currencies, the size of the market in any given country will depend upon the range of transactions subject to exchange controls, as well as the degree to which these restrictions are enforced by the authorities. In countries where demand rationing is not intense in the official market for foreign exchange, the parallel market will play only a marginal role. Conversely, in countries where balance-of-payments deficits are chronic and the central bank does not have sufficient reserves or borrowing capacity to satisfy the demand for foreign currency at the official exchange rate, parallel markets will typically be well developed and organized, with an exchange rate substantially more depreciated than the official rate.

The coexistence of an official and parallel market for foreign exchange results from the possibility of potential penalties, or, in other words, expected costs, imposed on private agents who fail to obey pricing or

---

5 The imposition of a tariff by itself creates incentives for smuggling but does not create incentives for the emergence of a parallel currency market. Such a market will usually emerge only if exchange controls are in place. In the particular case where legal trade requires the sale or purchase of legal foreign exchange, however, the existence of a positive tariff will also be sufficient to induce illegal trade and currency transactions (Pitt, 1984).
other regulatory directives (for example, surrender requirements). Pitt (1984) and Jones and Roemer (1987), for instance, suggest that the coexistence of legal and illegal markets depends on the way penalties are levied, that is, on the likelihood of getting caught. Both legal and parallel markets will exist if the risk of penalties can be reduced by engaging in legal sales that mask profitable but illegal transactions. Even if the risk of detection attaches only to illegal sales, however, the official market may still exist. This will occur if the penalties for illegal transactions drive the expected net marginal revenue from parallel-market sales below the official selling price. Without these requirements, and given the pressure of competitive forces, the parallel market is likely to collapse and a unified official market to emerge.

Parallel currency markets, although illegal, are often tolerated by the authorities in developing countries (for example, Bangladesh in 1972, the Dominican Republic in 1982, and Guyana in 1987). Although exchange dealers do not always advertise their services, “local” markets are substantially unified and the prevailing price is known to all who have an interest in it. In some countries, market users go through personal intermediaries, which may be why the market seems so uniform. In other countries, the market is dominated by a small number of “big” operators, who fix the exchange rate, sometimes on a daily basis, using their judgments about supply and demand. They are followed by a large number of intermediaries who are physically present in the market on a daily basis. The spread between the rate the intermediaries pay and the rate the major operators pay them is the source of the intermediaries’ income and is reflected in the emergence of a spread between asking and trading rates. One consequence of this type of intermediation is that the actual size of the market is difficult to evaluate, and estimates are subject to wide margins of error.

**Supply and Demand for Foreign Currency**

Transactions in parallel currency markets usually take place in cash, but checks are also commonly used in some countries. In markets where the risk of default is low and the surveillance of international

---

8 Greenwood and Kimbrough (1986) rationalize the existence of a parallel currency market with a cash-in-advance requirement that forces individuals to accumulate foreign currency (either officially or illegally) before they can consume.

9 This does not preclude substantial variations within countries. In Guyana, for instance, the exchange rates offered in border towns are significantly more depreciated than those quoted in the “Wall Street” area of Georgetown (Thomas, 1989).
transfers ineffective, transactions in foreign currency notes are sometimes completed abroad. In Latin America and Asia, the principal traded items are U.S. currency notes, although bilateral trade with the United States accounts for only a small share of some of these countries’ external transactions. Sources of supply and demand vary from country to country and depend heavily on the nature and effectiveness of exchange restrictions imposed by the authorities.

The supply of illegal foreign currency comes in general from five possible sources: smuggling of exports, under-invoicing of exports, over-invoicing of imports, foreign tourists, and diversion of remittances through nonofficial channels. Government officials may also allow diversion of foreign exchange from the official to the parallel market in return for bribes and favors. Although all of these sources are likely to be used to some degree in most circumstances, there is in general a dominant source at each time and in each country. The smuggling of exports, for instance, is considered to have been a major source of supply in Pakistan, India, and Turkey in the early 1970s (Gupta, 1981, 1984). More recently, from 1977 to 1983, under-invoicing of exports as a percentage of official exports is judged to have been 20 percent for Argentina, 13 percent for Brazil, and 34 percent for Mexico (Gulati, 1988). Foreign tourism is regarded as a dominant source of supply in the Caribbean countries. Worker remittances have represented the key component in Egypt (Bruton, 1983), Morocco, Turkey, and Sudan, and in Pakistan in the late 1970s. For Pakistan, Banuri (1989) estimates the volume of illegal remittances to have been anywhere between 15 and 35 percent of the officially recorded amount. This source alone of illegal dollars amounted to 20 to 47 percent of international reserves (excluding gold) in 1983, and 8 to 20 percent of the official money stock—a quite significant increase in liquidity. In the case of Bangladesh, studies in the early 1980s found that 35 percent of the migrants remitted their savings through private, informal channels. Similar observations have been made for several other remittance countries (Keely and Tran, 1989).

Remittances and tourism differ from commercial sources of foreign currency in that they necessitate no additional illegal transaction (Banuri, 1989). Smuggling of exports, by contrast, requires illegal

10 This may reflect the convenience of using the U.S. dollar in international transactions or a safe-haven effect. It may also reflect the importance of non-trade-related sources of supply and demand for foreign exchange in the parallel market.
transportation across the country’s borders. This raises the costs of supply because of the need for payoffs to officials and the risks of confiscation and other legal penalties. The parallel-market premium should therefore be high enough to compensate the suppliers for their real costs and risks. Unless there are significant economies of scale and learning by doing in smuggling activity, this argument also suggests that, all else being equal, the parallel-market premium will be lower when remittances constitute the major source of supply.

Available estimates, although generally subject to error, stress the importance of smuggling,\textsuperscript{11} under-invoicing of exports, and over-invoicing of imports as the major sources of supply of foreign currency in most developing countries.\textsuperscript{12} It should be noted, however, that the incentive for over-invoicing of imports exists only when the tariff rate on imported goods is sufficiently lower than the parallel-market premium. In a country with high tariff barriers, the price incentive is for under-invoicing (smuggling in) of imports rather than for over-invoicing—the one exception being, of course, the case of capital-goods imports, for which tariffs are generally lower than average, or even zero. It consequently appears likely that under-invoicing of exports is the single major source of unofficial currency supply from illegal trade. When there is a tariff on exports, under-invoicing allows the exporter to avoid the tariff and to sell the illegally acquired foreign exchange at a premium; when there is a subsidy on exports that is lower than the parallel-market premium, the sale of foreign exchange in the parallel market more than compensates for the loss of the subsidy. Thus, for given rates, the higher the parallel-market premium, the higher the propensity to under-invoice exports.

\textsuperscript{11} Smuggling may take place in legal or prohibited goods. Cocaine exports, for instance, are considered to account for a large share of the unofficial inflow of U.S. dollars into some Latin American countries. In Brazil, illegal trade (gold and coffee exports, in particular) is believed to account currently for nearly 30 percent of the foreign-currency supply in the parallel market (Novaes, 1990).

\textsuperscript{12} The extent to which traders engage in fake invoicing is typically measured using partner-country trade data. To investigate the scale of under-invoicing or over-invoicing of exports, for instance, one can compare exports to major partner countries, as shown by domestic data, with the corresponding imports recorded by the partner country. When the latter are larger, the evidence points to under-invoicing of exports. When making these partner-country comparisons, however, it is important to adjust the trade data for transport costs, timing of transactions, and classification of transactions. See McDonald (1985), Gulati (1988), and Arslan and van Wijnbergen (1989) for recent attempts to use these procedures to estimate the degree of under- and over-invoicing in trade transactions.
The demand for foreign currency in the parallel market reflects generally three activities: legal and illegal imports, portfolio diversification and capital flight, and residents’ travel abroad. The demand for foreign currency to finance legal imports stems from the existence of rationing in the official market for foreign exchange. The demand to finance illegal imports is for goods that are either prohibited or highly taxed and are smuggled into the country. The inherent confidentiality of transactions in the parallel market and the absence of legal accountability for anyone operating in it provide incentives to agents to use the parallel market for concealing illicit activities.

The portfolio motive is particularly acute in high-inflation economies and in countries where considerable uncertainty over economic policies prevails, because foreign-currency holdings represent an efficient hedge against bursts of domestic inflation. Econometric evidence suggests that, in middle-income developing countries, portfolio diversification represents a key determinant of the demand for foreign exchange in the parallel market (Agénor, 1991). Uncertainty about future inflation encourages a high degree of substitution between domestic and foreign currencies, which produces problems of monetary control (discussed below). Portfolio diversification may also take place through the parallel market for foreign exchange when countries impose restrictions on private-capital outflows. Attempts at circumventing the regulations are funded through the parallel market.

Welfare Implications of Parallel Markets

The existence of a parallel market in foreign currencies has important welfare implications. Conceptually, a sensible approach to analyzing those implications is to identify the welfare effects of exchange and trade controls and then evaluate the marginal welfare impact of parallel-market activities emerging from the existence of these controls.

Exchange and trade restrictions in developing countries have often been introduced in an attempt to defend an otherwise overvalued fixed exchange rate, to impose balance-of-payments adjustment in economies faced with limited foreign reserves and an external borrowing constraint (and, therefore, limited ability to defend the declared parity), and to insulate commercial transactions from the “disruptive” effects of transitory financial shocks. This has been particularly true in Latin America (Dornbusch, 1986).

Various rationales, emphasizing short- to long-term goals, have been put forward by policymakers to defend support of the exchange rate. Newly independent countries have often viewed an overvalued exchange
rate as a symbol of economic independence, with little concern (at least initially) about the economic costs of such a choice. In other cases, overvaluation has been perceived as an inexpensive way to provide cheap imports to domestic producers and consumers—notably imports of capital goods, durable goods, and intermediate inputs not produced domestically. Inexpensive access to such imports has been viewed as essential in promoting economic growth in the medium and long term. In some countries, an overvalued exchange rate has been viewed as a short-term anti-inflation device; by helping to keep down the domestic price of imported goods, an overvalued rate was believed to limit pass-through effects of changes in world market prices on domestic inflation. Another rationale often provided for adhering to an overvalued exchange rate relates to the perception that such a rate is capable of fostering the redistribution of income and economic activity from the tradable-goods sector to the nontradable-goods sector. Such a motive has often been an essential element of "inward-oriented" trade strategies of populist governments concerned about the concentration of wealth in the tradable-goods sector. Finally, exchange controls have been imposed in countries, again predominantly in Latin America, where short-term capital flows were perceived to be erratic and thought to lead to adverse movements in the real exchange rate (with negative consequences for the current account) or to produce recurrent speculative attacks that would exhaust the authorities' stock of foreign reserves and hamper financing of trade transactions.

Whatever the rationale for the imposition of exchange controls, however, there are welfare costs associated with the restrictions. Some of these costs have been well documented in the literature on international trade. An essential result, emphasized by Bhagwati (1978) and Greenwood and Kimbrough (1986), is that exchange controls place a quota on imports, thus raising their domestic relative prices just as tariffs would. To the extent that exchange controls lead to the emergence of a parallel currency market, however, they also affect private agents' economic decisions—notably the decision to evade restrictions by purchasing foreign currency illegally in the parallel market. The additional welfare effects are complex, because they vary according to the category of agents considered, and they have not yet been fully examined in the literature.

The gains and losses associated with parallel markets depend on a number of factors, in particular the penalty structure. If the expected costs of using these markets are low for private transactors, sellers as well as buyers, welfare is likely to be higher than if only official channels
are used. For instance, workers abroad remitting funds and foreign tourists selling dollars will get more units of domestic currency at the parallel rate than at the official rate. If penalties (fines, prison terms, etc.) are enforced to some degree, however, expected costs may be quite high. It is not possible, in general, to quantify the exact magnitude of gains and losses, but it can be shown that, in the case of smuggling, losses to smugglers can outweigh gains to consumers (Bhagwati and Hansen, 1973). This occurs when smuggling operations are subject to rising costs because of penalties, so that illegal imports replace official imports without lowering the cost of imports to domestic consumers. This result need not carry over to other forms of cheating, however, such as fake invoicing and diversion of remittances. If it is not costly to manipulate invoices, welfare gains to exporters will probably outweigh potential losses (Gupta, 1984).

From the point of view of the authorities, parallel markets have some obvious adverse effects. First, there is a cost of enforcement to counteract illegal activities and punish offenders. Second, there is a loss of tariff revenue as a result of smuggling and under-invoicing, a loss of income taxes and domestic indirect taxes, and a reduced flow of foreign exchange to the central bank, which lowers the government’s capacity to import. Third, parallel markets encourage rent-seeking activities (corruption of government officials, for instance), which lead to a suboptimal allocation of scarce resources. Fourth, the existence of a parallel market facilitates the switch from domestic-currency assets to foreign-currency assets (foreign-currency balances held domestically or interest-bearing assets held abroad) and may reduce the seigniorage revenue accruing to the government.13

Despite these costs, parallel markets are widely tolerated in developing countries. The typical argument used to justify them is that governments realize that, as long as there is demand rationing in the official market for foreign exchange, there is bound to be a secondary market, which can be eliminated only at prohibitive cost. Viewed in this way, a parallel market in foreign currency is taken to be socially desirable—even though the authorities’ ultimate goal is to remove discriminatory practices and stress legality in economic activities—because the parallel market meets the demands of operators rationed in the official market.

13 Whether or not seigniorage revenue falls depends on the type of exchange arrangement in place. In a crawling-peg regime, the authorities may be able to compensate for the loss of seigniorage by setting the rate of crawl—and therefore the rate of inflation, which determines the rate of return on domestic-currency assets—to generate a given level of revenue. Such an option is precluded in a fixed-rate regime.
McDermott (1989) puts forward another interesting argument that may help explain why authorities tend to accommodate rather than confront parallel markets. He suggests that the existence of a parallel currency market may yield two types of benefits. First, it increases employment by raising the domestic availability of imported inputs. Second, it may actually increase the flow of foreign currency to the central bank. This latter, somewhat paradoxical, effect may arise when the increased availability of imported inputs allows total exports to expand and to expand so much that foreign-currency receipts increase through both legal and illegal channels. There is, however, little empirical evidence to support this view. Overall, the welfare effects of exchange controls in the presence of parallel markets are largely ambiguous at the aggregate level.

3 Determinants of Parallel Exchange Rates: Theory

Over the past few years, parallel markets for foreign exchange have been analyzed and modeled from a number of different perspectives. In this section, we first examine “real trade” models of parallel markets and then focus on the portfolio-balance approach, which has recently attracted considerable interest.

Following the early partial-equilibrium analyses by Boulding (1947) and Michaely (1954) of a market for consumption commodity subject to price control and rationing, real trade models of the determination of the parallel-market premium focus solely on the parallel market itself and neglect its interactions with the rest of the economy. Specifically, the parallel market for foreign exchange is modeled as reflecting the demand for foreign currency to purchase illegal imports and the supply of foreign currency derived from illegal sources. Martin and Panagariya (1984), McDermott (1989), Sheikh (1976), and Pitt (1984) emphasize the role of smuggling and under-invoicing of exports as the main sources of supply, whereas Culbertson (1975) stresses the resale of officially allocated foreign exchange. This class of models emphasizes the impact of high trade taxes on smuggling activities and illegal currency transactions. As shown by de Macedo (1987) and Branson and de Macedo (1989), an importer will tend to smuggle if the tariff is so high that it pays to purchase foreign exchange at a premium in the

---

14 A detailed description of alternative analytical models of parallel currency markets (which include, in addition to those discussed here, the monetary model and models of formal dual-exchange systems with leakages) is provided in a previous version of this essay, available from the author upon request.
parallel market, even after allowing for the possibility of getting caught by the customs enforcement agency. If $\tau$ denotes the tariff, $\pi$ the probability of success in smuggling, and $\rho$ the premium, a necessary condition for import smuggling to occur is $\pi \tau > \rho$. Similarly, under the same detection technology, an incentive to smuggle out exports will exist when $\hat{\theta} < \pi \rho$, that is, when the subsidy (or tax rate) on exports $\hat{\theta}$ is smaller than the parallel-market premium weighted by the probability of success in smuggling.

In this framework, planned smuggled imports provide the flow demand for foreign currency in the parallel market while successfully smuggled exports provide the flow supply. The long-run parallel-market premium is then determined by the equilibrium conditions for legal and illegal trade. In the long-run equilibrium, where legal exports equal legal imports and successfully smuggled exports pay for planned smuggled imports, the premium can be expressed as a weighted average of $\tau$ and $\hat{\theta}$, and it is therefore determined (as is the smuggling ratio) by the structure of tariff barriers.

Real trade models provide an adequate framework for analyzing the impact of trade restrictions (as distinct from exchange controls) on the parallel-market rate. The basic limitation of their approach is that, because the only reason to deal in foreign currency is to buy imported goods, the sole purpose of black-market activity is to enable smuggling to take place. This assumes away the portfolio motive that has been identified as a critical contributor to the demand for foreign currency. Moreover, although this approach provides a useful analysis of the long-run determinants of the premium, it contains no mechanism that satisfactorily explains the short-run behavior of the premium, which is taken as given by exporters and importers in most models.\(^\text{15}\)

Attention has recently focused on the portfolio-balance approach, developed by de Macedo (1985, 1987) and Dornbusch et al. (1983), which stresses the role of asset composition in the determination of the parallel-market rate.\(^\text{16}\) Portfolio diversification has indeed been identified as a critical component of the unofficial demand for foreign currency in many developing countries.

\(^{15}\) The approach can be extended to do so, however. De Macedo (1987) develops a model in which the long-run premium is determined by the structure of trade taxes while the short-run premium results from the requirement of portfolio balance, as described below.

The general observation underlying this class of models is that foreign exchange is a financial asset. Loss of confidence in the domestic currency, fears about inflation and increasing taxation, and low real domestic interest rates give rise to a demand for foreign currency, both as a hedge and refuge for funds and as a means of acquiring and hoarding imports. In portfolio models, expectations play a key role in determining short-term shifts in supply and demand and in accounting for the volatility of parallel-market rates. An anticipated future change in the domestic money stock will begin to exert its effects immediately after announcement, for instance—or as soon as agents become aware that the policy change will take place—and will generate portfolio readjustments as well as concomitant changes in the parallel-market rate, so as to achieve the desired composition of private agents’ portfolios.

Although the partial-equilibrium formulation of Dornbusch et al. (1983) assumes the existence of domestic and foreign interest-bearing assets, the essential features of the approach are best captured by models in which domestic agents hold in their portfolios only non-interest-bearing domestic and foreign money. These models are based on the currency-substitution hypothesis, whereby money balances denominated in foreign currency are assumed to represent a substitute for domestic money as a store of value, unit of account, and medium of exchange. They provide considerable insight into the short- and long-run behavior of parallel-market exchange rates.

In all these models, output is exogenous, and the desired proportion between domestic and foreign currencies is given by a liquidity preference function (Calvo and Rodriguez, 1977) that depends on the expected—and, under perfect foresight, actual—rate of depreciation of the parallel-market exchange rate. Private capital transactions through the official market are usually ignored, so that the reported current-account balance is equal to the change in central-bank reserves, which, together with an exogenously determined rate of growth of domestic credit, determine the changes in the domestic money stock. The unreported current-account balance determines the change in the stock of foreign currency held in private agents’ portfolios. The flow supply of foreign exchange in the parallel market usually derives from under-invoicing of exports. The propensity to under-invoice, when endogenous, is assumed to depend positively on the level of the premium. The probability of detection is also assumed to rise as fraudulent transactions increase, and this translates into a rising—but at a diminishing rate—marginal under-invoicing share.

Portfolio balance implies that the domestic currency value of the
stock of foreign assets is equal at each instant to a desired proportion of private wealth. In the short run, the parallel-market rate moves so as to set the portfolio demand for foreign assets equal to the existing stock of foreign currency, implying that flow demand and supply may diverge at any given moment. The determination of the parallel exchange rate at any moment is thus made using the portfolio-balance equation, with the stock of foreign currency assumed to be fixed. In the long run, the parallel rate and private-sector holdings of foreign currency are determined jointly by the requirements of both portfolio and current-account equilibrium.

Although there remain important differences between individual formulations, some general conclusions can be derived from this class of models. Under a fixed-rate regime, an expansionary fiscal and credit policy generates a depreciation of the parallel exchange rate, a rise in prices, a real appreciation of the official exchange rate, and a decline in the prices of goods for which export proceeds are surrendered through the official market relative to the parallel market. As a consequence, the proportion of export proceeds repatriated at the official exchange rate falls, and official reserves decline. Eventually, the central bank will run out of reserves, and a balance-of-payments crisis will ensue. At this point, the inconsistency between expansionary macroeconomic policies and a pegged official exchange rate will become unsustainable, and corrective measures will need to be implemented—in the form of a parity change, for example. The process leading to a devaluation crisis has been well documented by Edwards (1989) and Edwards and Montiel (1989). The macroeconomic effects of a devaluation in this class of models are examined below.

4 Policy Issues with Parallel Currency Markets

A variety of policy problems are posed by the existence of parallel currency markets. The analysis here will focus on three key macroeconomic issues: the effectiveness of exchange restrictions, the use of

---

17 For instance, Edwards and Montiel (1989) consider a three-good economy and develop a fairly general analytical framework, but they assume that foreign-currency holdings remain constant, thus excluding an important source of dynamics.

18 In addition to its impact on the propensity to under-invoice exports, an increase in the premium, without an equivalent increase in domestic prices, may generate a positive wealth effect on aggregate demand, which may further worsen the current account of the balance of payments.
nominal devaluations to control the parallel-market premium, and strategies for unifying official and parallel markets.

Effectiveness of Exchange Controls

Many objectives have been put forward to justify the imposition of exchange controls in developing countries. Some have been discussed above. The evidence suggests, however, that controls have not been very effective in attaining any of these objectives. Rationing has created shortages of imported goods and shortages have encouraged smuggling. This has made it highly lucrative for rent-seeking traders to resell in the parallel market foreign exchange illegally acquired at the official rate. The existence of a high and positive premium in the parallel market has represented in many countries a strong incentive to divert export receipts from the official to the parallel market. As a result, instead of increasing the reserves at the disposal of the authorities, official controls have succeeded only in diverting a substantial part of the available foreign exchange to illicit use.

High parallel-market premia have considerably weakened the balance of payments in some countries, with potentially dramatic effects. If agents with forward-looking expectations are aware of a limit on the level of reserves the central bank can commit in defense of an officially fixed exchange rate, they may accelerate the diversion of export receipts from the official to the parallel market, which may precipitate a balance-of-payments crisis and the collapse of the fixed rate (Agénor and Delbecque, 1991).

The expansion of a parallel market for foreign exchange also weakens the effectiveness of capital controls. Formally, it has effects similar to an increase in capital mobility—which may accelerate capital flight and thus lead to an increase in the degree of substitutability between domestic and foreign currencies. The potential for currency substitution becomes an effective way of avoiding the inflation tax on holdings of domestic cash balances. The shift from domestic to foreign money results in a loss of seigniorage for the government, which, for a given real fiscal deficit, may call for a higher inflation rate achieved by monetary expansion and recurrent devaluations of the official exchange rate or, under a crawling-peg regime, an increase in the rate of depreciation (Agénor, 1990b).

19 Kamin (1991b) has reported empirical evidence that recorded exports are inversely related to the size of the parallel-market premium for a large group of developing countries. See also Arslan and van Wijnbergen (1989).
Finally, the rationing of foreign exchange may have negative output and employment effects (Austin, 1989). By reducing the supply of intermediate goods to import-dependent industries, including the export-oriented sector, exchange controls may reduce both the official supply of foreign exchange as well as economic activity to low levels of capacity utilization. They may, therefore, aggravate the very problems they were intended to solve.

The logical and obvious implication of this discussion is that, if parallel markets emerge in response to the imposition of controls, the most effective remedy is to eliminate the restrictions and let prices reflect fully the scarcity of foreign exchange. Indeed, in the past decade, several developing countries have shifted toward relatively less restrictive trade and exchange regimes (Quirk et al., 1987, 1989).

Devaluations and Parallel-Market Premia

The existence of a parallel currency market in which transactions take place at an exchange rate more depreciated than the official exchange rate is typically considered prima facie evidence that the official parity is inappropriate. Under these circumstances, the questions of whether or not to adjust the official exchange rate and to what extent to adjust it pose major issues for exchange-rate policy.

The view that a once-and-for-all devaluation of the official exchange rate may permanently reduce the level of the premium is a recurrent theme in discussions of exchange-rate policy in developing countries. Indeed, an argument often made is that a devaluation will reduce the incentive to fake foreign-trade invoices if a sizable parallel market exists, thereby attracting foreign exchange back to the official market. This argument has been used on various occasions to justify attempts to reduce the parallel-market premium by devaluation. In October 1989, the USSR announced a new exchange rate for foreign travel of 6.26 rubles per U.S. dollar (compared to a previous 0.63 rubles) to stop the leak of hard currency through the parallel market. In November 1989, the Argentine government announced it would reduce the 54-percent premium in the parallel market through fiscal reforms, rather than devaluation, which would fuel inflation; the austral was devalued by 54 percent in early December 1989, however, with the premium being cited as one reason for the decision to adjust the

---

20 The active repression of parallel markets has been attempted by some countries (Guyana in 1980, Tanzania in 1983, Algeria in 1990). It has proved difficult, however, to maintain an aggressive or punitive stance against well-entrenched informal activities.
exchange rate. In April 1990, the Nicaraguan government devalued the cordoba by 30 percent against the U.S. dollar and presented the measure as a key instrument in its attempt to control black-market activities. And in June 1990, the authorities in Guyana devalued the domestic currency by 36 percent relative to the U.S. dollar and announced that a series of devaluations would take place in 1991 until the official and parallel market rates for the currency were equal. The difficulty with this strategy comes, of course, from its partial-equilibrium nature and its neglect of macroeconomic interactions. The currency-substitution models described in Section 3 provide a useful analytical framework for examining these issues.

Consider, first, a fixed-rate regime in which cross transactions exist between the official and parallel market, and the extent of under-invoicing depends endogenously on the premium (Kamin, 1991b). Under this regime, the long-run effect of a once-and-for-all official devaluation on the parallel rate is ambiguous. The effect will depend on the degree to which fraudulent transactions react to changes in the premium, the rationing scheme imposed by the central bank, and the elasticity of export volume with respect to changes in relative prices. The greater the response of under-invoicing to the change in the premium, the greater the central bank’s tendency to hoard foreign-exchange receipts, and the higher the response of export volume, the more likely it is that the parallel market rate will depreciate less than proportionally, so that the premium falls in response to a parity change.21

The short-run behavior of the parallel rate and premium in response to a devaluation will reflect the typical behavior of asset prices. Consider, first, the case in which the devaluation is unexpected. Because it reduces the premium at the initial parallel rate, the parity change causes a decline in the flow of foreign currency to the private sector, and a depreciation of the parallel rate is required for the current-account balance to be maintained.22 At the moment of the devaluation, then, the parallel rate depreciates sharply. Subsequently, current-

21 Nowak’s (1984) result, according to which an official devaluation will always be associated with an appreciation of the parallel exchange rate, depends critically on the assumption that the central bank does not accumulate foreign exchange (Kamin, 1991b).

22 A depreciation of the parallel rate, given the official rate, increases the share of exports channeled through the unofficial market by under-invoicing or smuggling and thus increases the flow supply of foreign exchange. Conversely, import demand will fall, as well as the share of imports channeled through the parallel market by over-invoicing or smuggling, which will in turn decrease the flow demand for foreign currency.
account losses of foreign currency drive the unofficial rate up still further, until it reaches a new long-run equilibrium at the same time that foreign-currency holdings reach their new equilibrium level.

Suppose now that the devaluation is anticipated, because it is announced before being implemented. The announcement raises immediately the anticipated rate of depreciation of the parallel rate, and the actual rate of depreciation as well, for expectations are rational. Hence, the parallel rate depreciates immediately, and foreign-currency holdings rise. After the initial jump, the parallel rate continues to depreciate while private agents accumulate foreign currency in their portfolios. The economy reaches a new equilibrium trajectory at the exact moment the devaluation occurs. From this point on, the parallel rate continues to depreciate, but foreign-currency holdings decline because the unofficial current account (defined as the difference between unrecorded exports and imports) deteriorates following the devaluation. On the date of the announcement of the future devaluation, under-invoicing jumps, and it grows as the parallel rate depreciates. When the devaluation is implemented, the premium and under-invoicing fall sharply; they recover partly thereafter, because the parallel-market rate continues to depreciate until it reaches its new steady-state level.

This description of the transmission of a parity change provides an interesting explanation for the seemingly puzzling empirical results of sixty devaluation episodes in developing countries described by Kamin (1988). Prior to the typical devaluation, the growth rates of exports and imports fall sharply, while the current-account balance and reserve levels deteriorate markedly. Immediately following the devaluation, exports recover strongly and the current account improves (contrary to what a J-curve model would predict), while imports continue to fall, albeit at a slower pace, rebounding sharply in the second year after the devaluation. Kamin (1991b) provides a rationale for this sequence. Continuous inflation and, hence, real appreciation of the official exchange rate lead to increases in the parallel-market premium, increases in export under-invoicing, and reductions in officially measured exports. This drop in export proceeds leads, in turn, to reserve losses and declines in imports, as the authorities tighten foreign-exchange allocations. The expectation that the deteriorating external balance will prompt an official devaluation induces a speculative rise in the parallel-market rate that further reinforces the need to adjust the official exchange rate. Following the devaluation, the parallel-market premium falls, reducing under-invoicing and increasing officially recorded exports.
Improved reserve flows allow the authorities to expand progressively their sales of foreign exchange, so imports increase as well.

Consider now the case of a crawling peg. The official and parallel exchange rates depreciate at the same rate in the steady state, leaving the spread unaffected. An increase in the rate of devaluation of the official exchange rate leads to an equivalent increase in the rate of depreciation of the parallel rate, which generates a portfolio shift away from domestic money holdings. If the official and parallel markets are effectively segmented, the supply of foreign currency to the parallel market is fixed in the steady state, and an increase in the premium is required to restore portfolio equilibrium. The increase in the steady-state level of the premium caused by a higher rate of official depreciation has been emphasized by Dornbusch (1986) and Pinto (1986). It is important to note that, in these models, the steady-state premium does not depend on the level of the official exchange rate but only on its rate of change. This implies that discrete, one-shot devaluations will reduce the premium only temporarily in the absence of fundamental changes in fiscal and monetary policies.

In more elaborate models, however, there exists a potential source of ambiguity concerning the long-run effects on the premium of an increase in the rate of crawl. It relates to the role of the exchange-rate differential as an implicit tax on exports (Pinto, 1989, 1991; Kharas and Pinto, 1989). On the one hand, a higher rate of devaluation raises the rate of depreciation of the parallel-market rate, making foreign-currency holdings more attractive. This, by itself, would raise the premium. On the other hand, the policymaker may be faced with a given real fiscal deficit, resulting, for instance, from an exogenous level of domestic output and “incompressible” expenditures. A smaller domestic currency base will then be required to generate a given amount of revenue from the inflation tax. This, by itself, has an ambiguous effect on the premium; whether the premium rises or falls depends on the degree to which inflation affects the share of domestic-currency holdings in total financial wealth. If the elasticity of the share is less than unity, raising the rate of devaluation of the official exchange rate raises the unit yield of the inflation tax and lowers the parallel-market premium. Otherwise, the premium will actually rise. The elasticity itself, however, rises with the rate of inflation; the propensity to shift

23 Because there are two prices at which foreign exchange can be bought and sold, exports for which the proceeds are surrendered at the official exchange rate are taxed relative to other exports.
into foreign currency to avoid the inflation tax becomes stronger as the inflation rate rises. This results in a “seigniorage Laffer curve,” with the unit yield of the inflation tax rising for inflation rates below the seigniorage-maximizing inflation rate and falling above it. A similar reasoning yields a U-shaped curve linking the steady-state premium and the inflation rate, representing the trade-off between the export tax and the inflation tax. Overall, therefore, the impact of a devaluation on the parallel-market premium is ambiguous, whether one considers an officially fixed exchange rate or a crawling peg.

These predictions of currency-substitution models have recently been subjected to formal empirical tests. The evidence generally supports the presumption that parallel-market rates depreciate, although less than proportionally, in response to a devaluation of the official exchange rate, and that the premium falls initially. The evidence also suggests, however, that the reduction in the premium will be only temporary if fiscal and credit policies maintain an expansionary course, thus implying that a devaluation cannot by itself permanently lower the premium. Studies by Edwards (1989), Edwards and Montiel (1989), and Kamin (1991b) covering a large sample of devaluation episodes in developing countries have documented these facts. Similarly, in the empirical model presented in Agénor (1990a), a once-and-for-all devaluation of the official exchange rate is associated in the short run with a less-than-proportional depreciation of the parallel rate; in the long run, the official devaluation results in a permanently higher price level and a more depreciated parallel rate, with no effect on the premium. The econometric results presented in Agénor (1991) also support the view that the parallel rate depreciates less than proportionally immediately after an official parity change.

Dornbusch et al. (1983) present empirical tests of their model for Brazil, and Phylaktis (1991) considers the case of Chile. The results show a significant impact of the interest-rate differential—as well as the intensity of capital restrictions in the Chilean case—on the parallel-market premium. Fishelson (1988), using the actual rate of depreciation of the parallel rate as a proxy for the expected rate of devaluation of the official rate, tests the model provided by Dornbusch et al. for a group of nineteen countries for the period from 1970 to 1979. More recently, Kaufman and O’Connell (1990) have examined the case of Tanzania for the period from 1967 to 1988. Portfolio factors are shown in their analysis to affect the short-term behavior of the parallel-market premium; “real” factors play a predominant role in the long run.

The recent experience of Argentina also illustrates these propositions. Following the devaluation of December 1989, the premium dropped immediately. In the absence of financial discipline, however, the free-market rate rose quickly to 1,230 australes, bringing the premium back to 23 percent. See Kamin (1991a) for a further analysis.
There is evidence, moreover, that devaluations aimed at maintaining the premium below a given level may lead to increasing rates of depreciation and therefore to accelerating inflation. The Bolivian experience in the early 1980s is an often-cited example (Kharas and Pinto, 1989). In the three years preceding 1983, the inflation rate in Bolivia reached 47, 32, and 124 percent, respectively. In 1983, it reached 276 percent. In the last quarter of 1984, with the premium at 174 percent (compared to 22 percent in December 1980), there was considerable pressure to unify exchange rates. The authorities decided to reduce the spread by devaluing the official rate. They acted on the belief that the equilibrium nominal exchange rate is a weighted average of the official and parallel rates and that the official rate should be devalued toward the parallel rate. This resulted in a rule that directly linked official depreciation to the parallel-market premium. Official devaluation reached 350 percent in the last quarter of 1984 and 455 percent in the first quarter of 1985. The premium fell at first, but, with the parallel-market rate responding, it rose again, resulting eventually in an inflation rate of 496 percent in the first quarter of 1985, an annualized rate of 126,000 percent. This episode shows that targeting the premium by exchange-rate policy alone can be costly. Such a policy carries an inherent risk of inflation that can be limited only if highly restrictive financial policies are implemented at the same time.

Unification of Foreign-Exchange Markets

The unification of foreign-exchange markets, whereby the premium is lowered and the official and parallel rates are gradually brought closer so as eventually to produce a unique exchange rate, remains a key policy issue for many developing countries. These countries must also decide whether the unified rate should float or be pegged. When a parallel market is a significant source of import financing, the purpose of unification is to absorb and legalize it, rendering official the de facto

---

26 Bolivia has since moved to a fairly flexible exchange system based on daily auctions of predetermined amounts of foreign exchange without restrictions on access. Other countries recently pursuing an exchange-rate policy involving adjustment of the official rate to the parallel-market rate include Bangladesh and Ghana.

27 Policy discussions have often been centered on the idea that the restriction-free equilibrium exchange rate lies somewhere between the official rate and the parallel rate, although it has long been recognized that the latter is often subject to erratic movements due to fluctuations in the demand for, and supply of, foreign currency. In fact, as shown by Lizondo (1987), the equilibrium exchange rate can be either above or below the parallel rate.
import liberalization afforded by access to the parallel market and eliminating the inefficiencies and market fragmentation associated with it. Unification attempts may aim at adopting a uniform floating exchange rate, or a uniform fixed or crawling official rate. In the first case, the official exchange rate clears the foreign-exchange market; in the second, changes in the banking system’s foreign reserves serve to equate the supply and demand for foreign currency.

Consider, first, the policy of adopting a floating exchange rate. In the context of the theoretical framework described above, such a policy shift has ambiguous effects on the short- and long-run behavior of the exchange rate and the inflation rate. In the long run, the effect depends crucially on the fiscal impact of the exchange-rate reform. If the dual arrangement provides profits to the authorities in, for example, the form of tax revenues from currency operations, the rate of depreciation of the exchange rate and the inflation rate can be expected to rise, because the authorities are apt to compensate for a fall in revenue by an increase in monetary financing; conversely, if the system causes losses, the rate of depreciation and the inflation rate can be expected to fall.28

In the short run, the behavior of the floating exchange rate after unification will depend on a number of factors, in particular the behavior of expectations regarding the reform process. When the unification attempt is anticipated, agents seeking to avoid capital losses and to realize capital gains will shift immediately into foreign-currency assets if the uniform floating exchange rate is expected to depreciate relative to the existing parallel rate. They will shift into domestic-currency assets if the rate is expected to appreciate relative to the parallel-market rate. As a result of these portfolio adjustments, the parallel-market rate will move immediately—as soon as expectations are formed—toward the level asset holders expect the post-unification floating rate to reach. In the limiting case in which private agents anticipate perfectly the evolution of the post-unification exchange rate, the parallel-market rate will jump initially and then depreciate steadily toward that level as the time of actual unification approaches (Agénor and Flood, 1992; Lizondo, 1987; Kiguel and Lizondo, 1990).

28 The effect will also depend on whether the balance of payments is in deficit or in surplus before the unification attempt. For instance, an initial deficit, which implies that the excess demand for foreign exchange was partly accommodated through changes in international reserves, will translate upon unification into a higher rate of depreciation of the official rate and a higher inflation rate.
Consider now the case in which the authorities attempt to unify the official and parallel markets by adopting a crawling peg, possibly following a one-shot devaluation of the official rate. In the long run, the rate of crawl must be consistent with balance-of-payments equilibrium, and such a rate must equal the rate of depreciation that would prevail in the long run under a uniform floating regime (Lizondo, 1987). In the short run, the behavior of the parallel rate after unification will depend, once again, on the behavior of expectations. If agents anticipate the unification attempt, the portfolio adjustments described above will be initiated, and the parallel-market rate will move toward the expected level of the post-unification official rate. This illustrates the difficulty in using the parallel-market rate as an indicator for the initial level of the official rate under a crawling-peg regime. If private agents anticipate the unification attempt, the parallel rate will move immediately, before the reform is implemented, toward the level at which the authorities are expected to set the official crawling rate. Setting the initial, post-reform rate equal to the parallel rate at the time of unification will thus be consistent with balance-of-payments equilibrium only if expectations are correct.

What evidence is available concerning the behavior of exchange rates following a unification attempt? Few developing countries have attempted to unify their foreign-exchange markets by adopting a crawling peg. Furthermore, as argued above, a once-and-for-all devaluation cannot by itself lead to permanent reunification of the exchange markets. Hence, a sensible approach is to examine unification attempts that have taken the form of floating the domestic currency.

Recent evidence points to greater flexibility in the exchange arrangements of some developing economies, with several countries adopting market-oriented exchange systems. Roberts (1989) has studied the experience of African countries with market-based exchange-rate arrangements in the mid-1980s, specifically with foreign-exchange auctions and floating rates. These reforms often had as an explicit

29 Although a once-and-for-all nominal devaluation can be expected to reduce the premium, the reduction will be only temporary if fiscal and credit policies remain expansionary. Permanent unification of the official and parallel markets thus cannot be achieved by attempting to eliminate the spread solely by devaluation of the official rate.

30 Other developing countries recently adopting a floating rate include Uruguay in late 1982, Jamaica and the Philippines in 1984, and Bolivia and the Dominican Republic in 1985. The moves occurred in most cases at times of increasing external-payments difficulties, increasing arrears and capital flight (with reserves no longer available to support the fixed exchange rate), and with extensive parallel currency markets syphoning foreign
goal the absorption of the parallel market and a reduction or elimination of the premium. Table 1 summarizes the results. The rates of nominal devaluations that followed the introduction of auctioning or floating were massive in Nigeria, Sierra Leone, Somalia, Uganda, Zaire, and Zambia. The table shows clearly that the failures of auctions and floats are associated with a loss of control over monetary policy (for example, in Zambia and Ghana), whereas the successes are associated with a stabilization, if not a reduction, of liquidity growth (for example, in Gambia). Figure 2 presents monthly data on the spread for Guinea, Nigeria, Uganda, and Zaire. It shows that the parallel-market premium rose substantially before the reform of the exchange system. This rise can be interpreted as being partly the result of expectations about the reform process. The evidence also suggests that the premium fell sharply in all countries following the exchange-rate reform and that a significant premium reemerged subsequently in those countries where money growth could not be kept under control (Ghana, Sierra Leone, Somalia, Zambia).

Interestingly enough, the post-unification exchange rate is typically close to the pre-reform parallel rate, casting doubt on the argument that the “equilibrium” exchange rate is an average of the official and parallel rates. This should not be surprising when the resale of foreign exchange occurs on a large scale (as in Nigeria) and when it is remembered that prices are determined at the margin. A second misconception, pointed out by Pinto (1989), is that the large one-shot depreciation of the official exchange rate typically associated with unification must be inflationary. This has not been the case in countries where money growth was initially kept under control (Nigeria, Zaire), because the more depreciated parallel rate is already reflected in domestic prices. Inflation after unification seems to depend on the fiscal implications of unification and subsequent or concomitant changes in macroeconomic policies.

Fiscal factors can indeed account for a substantial rise in the infla-

---

31 As noted earlier, asset holders facing the possibility of a future depreciation of the parallel rate will reallocate their portfolios away from domestic money, thereby causing the free rate to depreciate immediately and the premium to increase prior to the depreciation of the official rate. The pattern depicted in Figure 2 is consistent with the results reported by Kamin (1991b), Edwards (1989), and Edwards and Montiel (1989).
<table>
<thead>
<tr>
<th>Country</th>
<th>Start of Float or Auction</th>
<th>Rate before Depreciation</th>
<th>Rate Immediately after Depreciation</th>
<th>Initial Nominal Depreciation of Official Rate (%)</th>
<th>Additional Nominal Depreciation over 1st Year of Float (%)</th>
<th>Broad Money Growth in 1st Year (%)</th>
<th>Inflation in Rental Prices in 1st Year (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gambia</td>
<td>1/86</td>
<td>3.5</td>
<td>7.0</td>
<td>6.8</td>
<td>49.0</td>
<td>10.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Ghana</td>
<td>9/86</td>
<td>90.0</td>
<td>150.0</td>
<td>120.0</td>
<td>25.0</td>
<td>30.0</td>
<td>65.0</td>
</tr>
<tr>
<td>Guinea</td>
<td>1/86</td>
<td>270.0</td>
<td>420.0</td>
<td>340.0</td>
<td>20.0</td>
<td>15.0</td>
<td>75.0</td>
</tr>
<tr>
<td>Nigeria</td>
<td>9/86</td>
<td>1.6</td>
<td>4.0</td>
<td>4.8</td>
<td>66.0</td>
<td>-12.5</td>
<td>15.0</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>7/86</td>
<td>5.0</td>
<td>15.0</td>
<td>12.0</td>
<td>58.0</td>
<td>78.0</td>
<td>126.0</td>
</tr>
<tr>
<td>Somalia</td>
<td>1/85</td>
<td>26.0</td>
<td>90.0</td>
<td>90.0</td>
<td>71.0</td>
<td>22.0</td>
<td>81.0</td>
</tr>
<tr>
<td>Uganda</td>
<td>8/82</td>
<td>100.0</td>
<td>360.0</td>
<td>300.0</td>
<td>67.0</td>
<td>-2.8</td>
<td>42.0</td>
</tr>
<tr>
<td>Zaire</td>
<td>9/83</td>
<td>5.8</td>
<td>66.0</td>
<td>30.1</td>
<td>80.0</td>
<td>26.0</td>
<td>35.0</td>
</tr>
<tr>
<td>Zambia</td>
<td>9/85</td>
<td>2.2</td>
<td>3.9</td>
<td>5.0</td>
<td>56.0</td>
<td>29.0</td>
<td>70.0</td>
</tr>
</tbody>
</table>

*Source: Roberts, 1989, p. 130.*
FIGURE 2
PARALLEL-MARKET PREMIA IN UNIFICATION ATTEMPTS
(in percentages)

SOURCE: International Monetary Fund and World Currency Yearbook.
NOTE: Asterisks mark start of float or auction.
tion rate following a unification attempt. Table 1 shows that inflation rose substantially in Sierra Leone in the first year following the attempt to unify markets by floating. An explanation for the often-observed inflationary burst related to unification has been provided by Pinto (1989, 1991). In developing countries, the government is typically a net buyer of foreign exchange from the private sector. Because the parallel-market premium is an important implicit tax, there is a trade-off between the premium (tax on exports) and inflation (tax on domestic currency holdings) in financing the budget deficit. Unifying official and parallel exchange rates can therefore raise inflation substantially and permanently if the level of government spending remains constant in real terms. The loss in revenues from exports is replaced by an increase in monetary financing of the budget deficit and a higher tax on domestic cash balances.

Two major lessons can be drawn from the recent experience of African countries with exchange-rate reform. First, unification of the official and parallel markets by exchange-rate policy alone cannot succeed without fiscal discipline. Second, complete elimination of the premium requires the removal of all restrictions on capital and commercial transactions. This conclusion largely corroborates the analysis of real trade models, discussed earlier. The combined effect of measures designed to relax import-licensing schemes and administrative allocations of foreign exchange is to make import transactions market determined, subject only to the distortions attributable to tariffs. In Ghana, for example, the exchange and trade system was gradually liberalized from 1986 to 1989, concurrently with the process of exchange-rate reform. The import-licensing scheme was first streamlined, then liberalized, and finally abolished in early 1989, and other current transactions were progressively made eligible for funding through the foreign-exchange auction. As a result of these measures, only a few restrictions on current transactions, relating essentially to invisible transactions, remained in effect by mid-1989. By contrast, in the other countries previously considered—notably Nigeria and Zaire—the currency was floated only for commercial transactions, with capital controls retained for outward flows. The maintenance of restrictions on capital flows, coupled with the absence of adequate monetary and fiscal policies prevented a substantial fall of the parallel-market premium in these countries.

The analysis suggests, therefore, that the “best” approach to unification might be to relax foreign-exchange rationing gradually in the official market, starting with commercial transactions and accompanying the
liberalization with discrete devaluations—the pace of reform being set by the speed and credibility of fiscal adjustment (Pinto, 1989). Monetary policy and liquidity controls are important to the process for stabilizing expectations, prices, and parallel exchange rates. Expectations of inflation and further depreciation at times of expansionary credit policies, typically caused by the monetization of fiscal deficits, may exert a destabilizing influence on the price level. In this sense, unification is a complex process, perhaps requiring institutional changes as well as behavioral adjustments on the part of market participants.

5 Concluding Remarks

This paper has sought to discuss in a consistent and coherent framework recent theoretical and empirical developments in the analysis of parallel currency markets in developing countries. The policy implications have also been assessed, and the issue of exchange-market unification has been discussed in light of the recent experience of a group of developing countries with flexible exchange rates.

A number of substantive issues have not yet been adequately addressed in the recent literature. The transition costs associated with exchange-market unification, for instance, are not well understood, and issues related to the distributional effects, as well as the pace, of the unification process have received only scant attention. In addition, the criteria for choosing between a gradual and an instantaneous adjustment and the implications of these alternatives for fiscal and monetary policies, for inflation, and for the balance of payments have not been systematically addressed.

Another potentially fruitful line of inquiry relates to the role of the premium as a signaling device in the context of stabilization programs. Recent analytical models of devaluation crises have emphasized the premium’s role in determining the degree of credibility forward-looking agents attach to the official exchange rate and its consequent effect on devaluation expectations (Agénor, 1990b). An interesting extension of this framework might be to analyze the role of the premium in an economy subject to a variety of stochastic shocks, in which the premium would convey noisy information about the policy stance of the authorities, and to examine how expectations of the collapse of a stabilization program can become self-fulfilling when agents face a signal-extraction problem.

Finally, it would be worth studying the implications of alternative exchange-rate rules (such as a constant real-exchange-rate policy) for the behavior of the parallel-market premium. One should also pursue a
systematic analysis of the impact of external shocks, such as changes in the terms of trade (as in Edwards and Montiel, 1989, and Pinto, 1986) on the behavior of the premium.

Nevertheless, some general policy lessons clearly emerge from the current literature. Exchange and trade restrictions are largely ineffective in the long term when directed at maintaining an overvalued exchange rate or “imposing” balance-of-payments adjustment. In these circumstances, the emergence of parallel exchange markets is a normal outcome. Although socially beneficial in some respects, parallel markets generate a variety of costs. In particular, to the extent that they provide a channel for portfolio diversification, they facilitate evasion of the inflation tax on domestic cash balances and may help accelerate capital flight. Furthermore, unofficial exchange rates have a substantial impact on domestic prices and play an important role in the transmission of macroeconomic policies.

Unification of official and parallel markets cannot be achieved by relying exclusively on devaluation of the official exchange rate to eliminate spread. Such a policy may also be inflationary, and attempts at unification by floating the currency may be accompanied by a burst of inflation. This burst is not primarily the result of the depreciation of the official exchange rate, for domestic prices will have already reflected the more depreciated parallel rate. It is the result of losing the implicit tax on imports through unification. To be successful, that is, sustainable, unification must be accompanied by a relaxation of exchange restrictions and by supportive fiscal and monetary policies.

References


Dornbusch, Rudiger, Daniel V. Dantas, Clarice Pechman, Roberto Rocha, and Demetri Simoes, “The Black Market for Dollars in Brazil,” The Quarterly


Notice to Contributors

The International Finance Section publishes papers in four series: Essays in International Finance, Princeton Studies in International Finance, and Special Papers in International Economics contain new work not published elsewhere. Reprints in International Finance reproduce journal articles previously published by Princeton faculty members associated with the Section. The Section welcomes the submission of manuscripts for publication under the following guidelines:

Essays are meant to disseminate new views about international financial matters and should be accessible to well-informed nonspecialists as well as to professional economists. Technical terms, tables, and charts should be used sparingly; mathematics should be avoided.

Studies are devoted to new research on international finance, with preference given to empirical work. They should be comparable in originality and technical proficiency to papers published in leading economic journals. They should be of medium length, longer than a journal article but shorter than a book.

Special Papers are surveys of research on particular topics and should be suitable for use in undergraduate courses. They may be concerned with international trade as well as international finance. They should also be of medium length.

Manuscripts should be submitted in triplicate, typed single sided and double spaced throughout on 8½ by 11 white bond paper. Publication can be expedited if manuscripts are computer keyboarded in WordPerfect 5.1 or a compatible program. Additional instructions and a style guide are available from the Section.

How to Obtain Publications

The Section’s publications are distributed free of charge to college, university, and public libraries and to nongovernmental, nonprofit research institutions. Eligible institutions may ask to be placed on the Section’s permanent mailing list.

Individuals and institutions not qualifying for free distribution may receive all publications for the calendar year for a subscription fee of $35.00. Late subscribers will receive all back issues for the year during which they subscribe. Subscribers should notify the Section promptly of any change in address, giving the old address as well as the new.

Publications may be ordered individually, with payment made in advance. Essays and Reprints cost $6.50 each; Studies and Special Papers cost $9.00. An additional $1.25 should be sent for postage and handling within the United States, Canada, and Mexico; $1.50 should be added for surface delivery outside the region.

All payments must be made in U.S. dollars. Subscription fees and charges for single issues will be waived for organizations and individuals in countries where foreign-exchange regulations prohibit dollar payments.

Please address all correspondence, submissions, and orders to:

International Finance Section
Department of Economics, Fisher Hall
Princeton University
Princeton, New Jersey 08544-1021
List of Recent Publications

A complete list of publications may be obtained from the International Finance Section.

ESSAYS IN INTERNATIONAL FINANCE

156. Sebastian Edwards, *The Order of Liberalization of the External Sector in Developing Countries*. (December 1984)


163. Arminio Fraga, *German Reparations and Brazilian Debt: A Comparative Study*. (July 1986)


*Publications marked by an asterisk are out of print. They are available on demand in xerographic paperback or library-bound copies from University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, USA, or 30-32 Mortimer St., London, WIN 7RA, England. Microfilm of all Essays by year is available from University Microfilms. Photocopied sheets of out-of-print titles are available from the Section at $9.00 per Essay and $11.00 per Study or Special Paper, plus $1.25 for domestic ($1.50 for overseas) postage and handling.*


175. C. David Finch, *The IMF: The Record and the Prospect*. (September 1989)


183. Michael Bruno, *High Inflation and the Nominal Anchors of an Open Economy*. (June 1991)


**PRINCETON STUDIES IN INTERNATIONAL FINANCE**


60. Thorvaldur Gylfason, *Credit Policy and Economic Activity in Developing Countries with IMF Stabilization Programs*. (August 1987)
64. Jeffrey A. Frankel, *Obstacles to International Macroeconomic Policy Coordination*. (December 1988)

**SPECIAL PAPERS IN INTERNATIONAL ECONOMICS**


**REPRINTS IN INTERNATIONAL FINANCE**

The work of the International Finance Section is supported in part by the income of the Walker Foundation, established in memory of James Theodore Walker, Class of 1927. The offices of the Section, in Fisher Hall, were provided by a generous grant from Merrill Lynch & Company.