

Exchange Rate Unification: The Cuban Case

Augusto de la Torre and Alain Ize[Ⓞ]

December 2014

Abstract

This paper reviews from an international perspective the challenges faced by Cuba in unifying its exchange rate and compares various options to meet this objective. It argues in favor of a fast unification approach but cushioned during a transition period by lump-sum taxes and subsidies applied on an enterprise-by-enterprise basis. By allowing for relative price changes to operate in full from the start, the immediate unification would maximize efficiency gains. At the same time, by cushioning the Cuban economy from potentially large transitional pains—including fiscal revenue losses, productive dislocations, inflationary outbursts and distributional effects—the lump-sum taxes and subsidies (to be gradually phased out in line with a preannounced schedule) would ease the unification, thereby boosting the credibility of the reform. However, to ensure the viability of the scheme and the rapid materialization of the efficiency gains, important habitating reforms would be needed, particularly as regard the governance of state enterprises.

JEL classification codes: F310, F410, F430

Keywords: dual exchange rates, exchange rate unification, economies in transition, dollarization, currency unification

[Ⓞ] This paper was prepared for a series of workshops on Cuban economic change in comparative perspective organized by the Foreign Policy Latin America Initiative at the Brookings Institution and the University of Havana's Center for the Study of the Cuban Economy and the Center for the Study of the International Economy. A nearly identical version of this paper was published by Brookings in an edited volume (<http://www.brookings.edu/research/papers/2014/11/cuba-economic-change-comparative-perspective>). The authors work for the World Bank as, respectively, Regional Chief Economist (adelatorre@worldbank.org), and Senior Consultant (aize@worldbank.org). The paper benefitted from comments by Aquiles Almansi, Tito Cordella, Eduardo Fernández-Arias, Daniel Lederman, Sergio Schmukler, and other participants at an internal World Bank workshop (August 2013), and comments by Richard Feinberg, Andrea Gallina, Guillermo Perry, José Juan Ruiz, Alejandro Santos, Alberto Trejos, Juan Triana, Pavel Vidal, and other participants at a Brookings-University of Havana workshop held in Havana (September 2013). Magali Pinat provided excellent research assistance. The views in this paper are entirely those of the authors and do not necessarily represent the views of the World Bank, its executive directors, or the countries they represent.

1. Introduction

Since 2011, the Cuban authorities have placed exchange rate unification as one of their top policy priorities. Indeed the current dual exchange rate system—whereby a one-to-one exchange rate for the “convertible peso” coexists with a twenty four-to-one exchange rate for the “Cuban peso” (both against the US dollar)—introduces severe and pervasive distortions with costly consequences for resource allocation and the growth potential of the Cuban economy. At the same time, the unusually large (by international comparison) spread between the two exchange rates exacerbates the transition costs and thus constitutes one of the main reasons delaying their unification.¹

This paper reviews from an international perspective the challenges faced by Cuba in unifying its exchange rate system and compares various options to meet this objective. It argues in favor of an immediate unification cushioned by a system of lump-sum taxes and subsidies to be phased out during a preannounced transition period. By allowing for relative price changes to operate in full from the start, the immediate unification would maximize efficiency gains. At the same time, by cushioning the Cuban economy from potentially large transitional pains—including fiscal revenue losses, productive dislocations, inflationary outbursts and distributional effects—the lump-sum taxes and subsidies would ease the transition, thereby boosting policy credibility. These lump-sum taxes and subsidies would be set on an enterprise-by-enterprise basis so as to fully neutralize, initially, the windfall losses or gains that individual enterprises would otherwise make upon the unification of the exchange rate.

By replacing at the outset the taxes and subsidies *implicit* in the current exchange rate spread with *explicit lump-sum* taxes and subsidies of equivalent magnitude, the traumatic effects the unification would have on impact would be eliminated. Thereafter, however, both the enterprises and the government would have to adjust, albeit gradually, because the lump-sum taxes and subsidies would be phased out according to a preannounced timetable. Note that, unlike proportional taxes and subsidies, lump-sum taxes and subsidies are set in absolute value. Once set, they do not vary with the level of production or profits of an enterprise and, hence, they do not affect the incentives of enterprises to produce or invest (at the margin). However, to stimulate a positive supply response (i.e., ensure that the change in relative prices associated with the exchange rate unification leads to an early materialization of efficiency gains) and to further ease transition pains as lump-sum subsidies are phased out, it would be necessary to implement important complementary reforms, particularly as regards the governance and management of state-owned enterprises and the overall tax and subsidy system.

The rest of the paper is organized as follows. Section 2 briefly reviews the international experience on exchange rate unification. Section 3 focuses on what is special about Cuba and what this implies for policy. Section 4 compares and contrasts the pros and cons of four policy options. Sections 5 and 6 provide rough sketches of how the proposed

¹ For information on the economic situation in Cuba and the circumstances underpinning the exchange rate system, see Di Bella and Wolfe (2008), Orro (2008), Dreher (2009), Vidal (2012a, 2012b), Vidal and Pérez Villanueva (2013), and Feinberg (2013).

fiscally-cushioned big bang option would work for two key sectors of the Cuban economy, the foreign-managed tourism industry and the state-owned enterprises that import and produce for the local economy, respectively. Section 7 concludes by discussing key habilitating reforms (fiscal, monetary and public sector governance) and some sequencing issues. It also briefly addresses the related yet distinct issue of currency unification.

2. Some lessons from international experience

Understanding well what caused the original dislocation that led to a multiple exchange rate regime is the inescapable starting point. Typically, multiple rate systems emerge after a large adverse shock hits the economy, exerts substantial pressure on the foreign exchange market, and hence calls for a major depreciation of the equilibrium real exchange rate. The shock can be a *supply* shock, such as the deterioration in the terms of trade that severely weakens the external trade account or an increase in world interest rates that sharply raises the servicing costs of the country's external debt. Or it can be a *demand* shock, such as a surge in local demand for foreign assets (capital flight) triggered by financial repression and/or unsustainable macro policies. In the first case, the required exchange rate depreciation reflects the deterioration in the country's purchasing power; in the second case, it reflects the relative price change needed to reduce the country's demand for tradable goods (imports, in particular) so as to allow its citizens to transfer their assets abroad. To avoid a politically explosive fall in real wages and a rise in inflation due to an increase in the cost of imported inputs, governments may introduce a dual exchange rate regime. The more depreciated "free rate" (the market determined rate) is used for capital account transactions and "non-basic" imports, while the less depreciated "official rate" is used for "basic" imports and export proceeds subject to surrendering requirements.

In theory, a multiple exchange rate regime can help limit inflation, protect socially sensitive economic activities, channel resources to developmental priorities, and redistribute income progressively (including by avoiding rewarding the owners of foreign currency-denominated assets through devaluation-induced valuation gains). In practice, however, such a regime causes large efficiency losses. The exchange rate differential acts as a tax on exports (through surrendering requirements) and a subsidy on "basic" imports, which are detrimental to the country's exporting and import-substituting productive activities, thereby hindering economic growth and job creation. Differentiating between "basic" and "non-basic" imports can give rise to similarly severe resource misallocations and opacities.

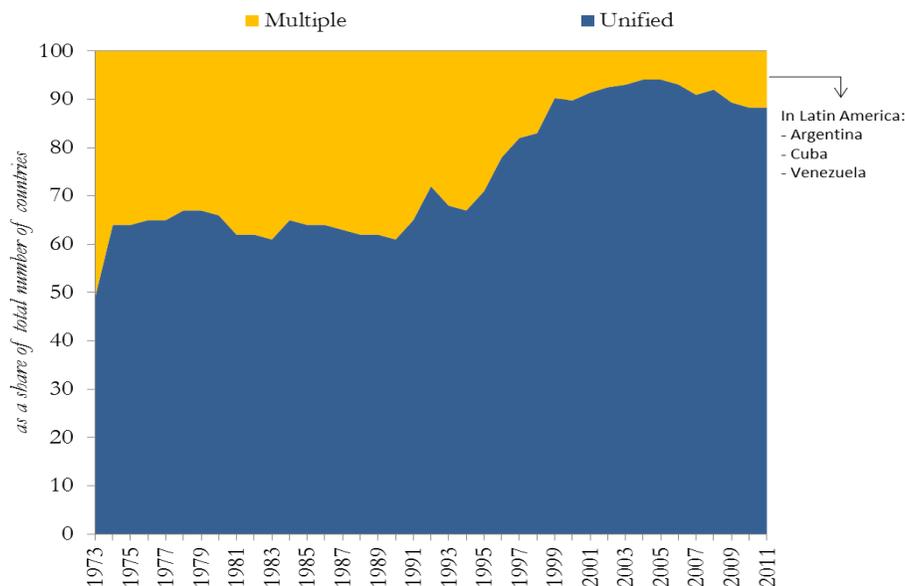
In all cases, the exchange rate spread causes growing wedges between private and social interests that translate into enforcement nightmares and multifaceted, socially destructive rent seeking activities. The costs of such distortions accumulate and worsen over time. As dual rates persist and become fossilized throughout the economy, they end up causing increasingly pernicious and ingrained segmentations between the winning sectors (that found access to preferential exchange rates) and the losing sectors (that did not).² Exchange rate unification should therefore give rise to both static efficiency gains

² Agenor (1992) provides a good summary review of the literature on parallel exchange rates, including causes, welfare impacts and policy implications. Frenkel and Razin (1986) analyze the restrictive conditions

(i.e., a better allocation of existing resources, so that more income and output can be generated with the already available labor, land, and capital) and dynamic efficiency gains (i.e., the income expansion arising from the process of capital accumulation and the more productive use of resources over time).

Reflecting the growing awareness of the efficiency costs of multiple exchange rates, but also the improvements in macro-monetary policy frameworks, the share of countries in the world with multiple exchange rate regimes across the world has steadily declined over the last forty years, albeit with a slight resurgence in the last 5 years (Figure 1). In Latin America, this resurgence included some churning between the return of parallel rates in countries that had successfully unified their exchange rates in the past (Venezuela and Argentina) and the recent unification of the exchange rate in countries with formerly dual exchange rate regimes, such as the Dominican Republic (Figure 2). Arguably, however, multiple exchange rates appear to be for the most part a species on the verge of extinction.

Figure 1: World Share of Multiple vs. Single Exchange Rate Regimes 1973-2011



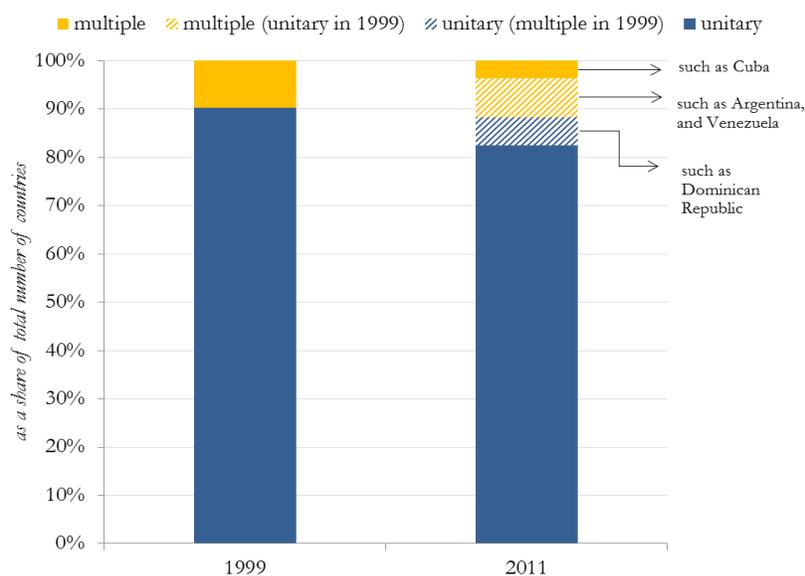
Sources: Rogoff et al. (2003) and IMF AREAER online database.

The persistence of multiple exchange rate regimes in the few remaining countries that still have them is likely to reflect in part the challenges and costs associated with exchange rate unification. These costs depend on the roots of the original split. In countries where the originating shock was a demand shock (e.g., massive capital flight) resulting from financial repression or poor macro-monetary management, a reversal towards sounder macro-financial policies may be all it takes to discourage capital flight and, hence, absorb the pressures that led the parallel rate to deviate from the official rate in the first

under which a dual exchange rate system that separates capital from current account transactions can be sustained indefinitely.

place. Instead, in countries where the shock came from the supply side and where the factors behind the shock still endure (for example a permanent worsening in the terms of trade), unification is likely to have more substantial costs.

**Figure 2: Churning Between Multiple vs. Unified Exchange Rate Regimes
1999-2011**



Notes: Countries with multiple exchange rates in 2011 but with a unified rate in 1999 are Angola, Argentina, Eritrea, Georgia, Guinea, Kyrgyz Republic, Malawi, Mongolia, Nigeria, Sao Tome and Principe, Sudan, Ukraine, Uzbekistan, and Venezuela. Countries with a unified rate in 2011 but with multiple rates in 1999 are Afghanistan, Belarus, Botswana, Cambodia, Dominican Republic, Egypt, Iran, Lao, Libya, Russia, and Turkmenistan. Source: IMF AREAER online database.

While the post-unification equilibrium exchange rate is often expected to lie somewhere within the two pre-unification exchange rates, in practice it ends up close to the pre-unification parallel rate. The reasoning behind this common expectation is straightforward. A depreciation of the more appreciated (official) exchange rate should reduce the demand for basic imports—thereby freeing some foreign exchange—or promote exports—thereby allowing additional foreign exchange to come in. As this net additional supply of foreign exchange finds its way into the market, it should allow the more depreciated (parallel) rate to strengthen. Thus, as the official exchange rate continues depreciating, the two rates should gradually move towards each other, eventually converging somewhere inside the initial spread. In practice, however, because the demand for basic imports is generally fairly inelastic and the response of exports to a more competitive exchange rate tends to take time to materialize (e.g., it is short-run inelastic), the depreciation of the official rate is unlikely to free much foreign exchange, at least in the short run. If so, unless there is a sufficiently deep, forward-looking foreign exchange market that anticipates future foreign exchange inflows, the two rates will tend to meet close to the bottom of the range, i.e., close to the parallel market rate.³ Moreover, there

³ This is what Agenor (1992) finds in the experience of exchange rate unifications in a number of African countries during the 80s. He concludes that “the post-unification exchange rate is typically close to the pre-

may be significant additional changes in the demand for foreign exchange coming from the capital account. For example, a surge in demand for dollars can materialize if there are expectations of a rise in inflation or if the central bank starts accumulating foreign reserves after the unification. Thus, the post-unification single exchange rate could well depreciate (overshoot) beyond the pre-unification parallel rate.⁴

Unless offset through fiscal adjustments (more on this below) a rise in the post-unification price level can generally be expected from the pass-through of the official exchange rate depreciation to the consumer price index (e.g., *cost-based inflation*).⁵ Moreover, the rise in the price *level* may turn into a permanent rise in the *rate* of inflation when the initial supply shock leads to wage-price spirals that are accommodated by a weak (or not credible) monetary policy (e.g., *expectations-based inflation*); or when the unification gives rise to lasting fiscal imbalances (e.g., *demand-based inflation*) resulting from the loss of quasi-fiscal income generated by the dual exchange rate regime (in particular, an abrupt reduction in the implicit taxation of exports).

Table 1: Post-Unification Inflation in Selected Latin American Countries

Country	Beginning of Unification Process	Pre-Unification Premium	Phasing	Point-to-Point Annual Inflation			
				3 months before unification begins	...3 months after	...12 months after	... 24 months after
Venezuela	Mar. 1989	202%	Fast	36%	103%	81%	36%
Argentina	Feb. 1989	53%	Gradual	372%	460%	8163%	232%*
Peru	Jun. 1989	166%	Gradual	3414%	5704%	1968%	207%*
Ecuador	Sept. 1992	27+%	Gradual	50%	64%	46%	27%
Dominican Republic	Oct. 2003	10+%	Fast	26%	40%	52%	0%
Cuba	?	2300%	?	4%	?	?	?

Notes: *Premium* is defined as the average spread between the parallel and official rates as a percent of the official rate in the last quarter before unification. *Phasing* is defined by the duration of the coexistence of official and parallel rates after unification—“fast” is less than 3 months. * 28 months after, to capture the effects of the formal introduction of Convertibility in April 1991 (Argentina) and the final currency unification in August 1991 (Peru). Sources: EIU, World Currency Yearbook (several publications), AREAER (several publications), Pick’s currency yearbook (several publications), Kiguel and O’Connell (1995), Marion (1999), Ilzetzki, Reinhart, and Rogoff (2008), Kamin (1991), IFS database.

reform parallel rate, casting doubts on the argument that the equilibrium exchange rate is an average of the official and parallel rates”.

⁴ See Lizondo (1987) and Agenor and Flood (1992).

⁵ However, such supply side pressures may be relatively subdued in countries where domestic prices already reflected for the most part the parallel rate (rather than the official rate).

Table 1 provides a bird-eye view of post-unification inflationary experiences in Latin American countries that have unified their exchange rate in the not too distant past. Inflation tended to be higher in countries with high pre-unification inflation (Argentina, Peru) or in countries with high pre-unification spreads (Venezuela, Peru). By contrast, for countries with relatively low pre-unification inflation (Ecuador, Dominican Republic), post-unification inflation was relatively subdued. For Cuba this is both good news and bad news. While pre-unification inflation is low in Cuba, the spread is way above anything observed in other countries of the region.

3. The case of Cuba

The Cuban dual exchange rate system overlaps with a dual currency system, featuring a Convertible Peso (CUC)—a fully convertible currency that exchanges at one to one against the dollar—and a Cuban Peso (CUP)—that exchanges at 24 to one dollar.⁶ Hence, the dual monetary or currency system (CUP and CUC) coexists with a dual exchange rate system (one-to-one and 24-to-one, in relation to the dollar). This paper concentrates on exchange rate unification issues and presents only a brief discussion of currency (monetary) unification issues in the final section. Note that the terminology used in the previous section—which distinguished between “official” and “parallel” exchange rates—does not apply to the Cuban case, where the two exchange rates are official.

The dual exchange rate system goes back to the early sixties when a separate exchange rate was used for trade with countries of the former Soviet Union. The foreign exchange market segmentation was abruptly exacerbated during 1990-1993, when political changes in the former Soviet Union led to a redefinition of commercial relations with Cuba that resulted in a huge worsening in Cuba’s terms of trade. The CUC was subsequently created (in 1994) to limit dollarization—by providing an alternative to the US dollar as a unit of account and a store of value—in the environment of rapidly rising and highly volatile inflation that followed the adverse terms of trade shock. This overlap and common origin has led many observers to see the dual exchange rate and the dual currency systems as “joined at the hip.” But they are clearly driven by different motives and pursue different objectives (more on this below). Exchange rate unification can thus be conceptually and practically de-linked from the currency unification.

What makes the case of Cuba special when compared to other cases of dual exchange rate regimes in Latin America? At least four features stand out. First, because the origin of the Cuban dual exchange rate system is primarily real (linked to a terms of trade shock) rather than financial, it basically amounts to a *quasi-fiscal scheme*. It taxes exports (including of course tourism) while subsidizing basic imports. This is both good and bad

⁶ Actually, the CUP exchanges at 24 CUP per dollar for households and the self-employed at 1 CUP per dollar for the state enterprises and institutions. Persons can exchange CUPs for dollars (at 24 to 1), as well as CUPs for CUCs (at 25 to 1), in the exchange bureaus (Cadecas). State enterprises and institutions conduct exchange transactions with the Central Bank and are prohibited from using the Cadecas to arbitrage between the exchange rates. There are no indications at this time, however, of a significant spread between the 24 per 1 USD rate at the Cadecas and the rate in informal transactions on the street. For details see Vidal (2012a).

news. It is good news because it would mitigate concerns about speculative financial turbulence during unification. Such concerns would be further limited by the fact that much of the population's savings are already held in the form of cash dollars (hence do not generate a pent-up demand for dollars) while most bank deposits, whether in CUPs or CUCs, belong to public enterprises (hence are unlikely to trigger speculative shifts into dollars). However, it is also bad news because, unless fiscally neutralized, the exchange rate unification would unavoidably have large re-distributional implications, making the former sellers of CUCs better off and the former purchasers (and their clients) worse off. While such re-distributions can be offset through bold yet temporary fiscal means (see below), the sustainability of the unification in the medium-term would require comprehensive fiscal reforms (more on this below).

Second, the spread between the two exchange rates in Cuba, at 2,300 percent, is by far the largest in post-World War II Latin American history. This implies a significant risk that the pass through effects of the depreciation of the CUC rate may unleash a wage-price spiral. To avoid it, a premium would need to be placed on maintaining a very tight control over monetary expansion during and after the unification.

Third, the Cuban the government can exert control over the actions of public enterprises arguably more than in other countries. This may complicate some things and facilitate others. The materialization of efficiency gains would require a sufficiently elastic supply response (in the form of decisions to increase investment and production) to the changes in relative prices that would result from unification. But this supply elasticity would depend on the quality of relative price signals and the responsiveness of decentralized economic actors to such signals (and to other market incentives), both of which are likely to be weak or sluggish in the large state sector of the economy. On the plus side, however, the control of public enterprises by the government can also help limit price increases on the most basic goods during the transition (more on this below).

Fourth, Cuba has limited access to international finance. This is an important complication. Indeed, concessional international finance, if available, could significantly facilitate unification by allowing expenditure to exceed income during the transition.

4. Exchange rate unification options

In view of the constraints mentioned above, *the key challenge for Cuba is to balance the short-term reallocation pains of the exchange rate unification with its medium and long-term efficiency gains.* Over time, by boosting the size of the cake, efficiency gains should provide plenty of room to offset the initial reallocation pain. Thus, beyond the transitional adjustment pains, unification should be a win-win for all. Indeed, the achievement of such efficiency gains in the functioning of the economy is the *raison d'être* of unification. The problem is that the size of the cake is largely given in the short-term, as the increase in the capital stock (including in foreign direct investment—FDI) and reallocation of resources that are required for efficiency gains would likely take time to materialize. Thus, unless cushioned in some way, the raw initial impacts of unification

(fiscal revenue losses, productive dislocations, inflation outbursts, and regressive distributional effects) could be quite painful.

A successful transition strategy should thus pursue two objectives. First, to limit the short-term pains until efficiency gains materialize. Second, to boost the pace at which such efficiency gains materialize. The extent to which these two objectives are fulfilled should provide the basic measuring rod with which to compare and rank available unification options. In the rest of this section, we distinguish four typological options.

Option One may be labeled *raw big bang*. It consists in unifying the exchange rates on day one. To limit balance of payments and foreign exchange market pressures, the exchange rate would be unified the rate of 24 (new) pesos per dollar. All exchange transactions would thereafter take place at the single new rate. The main advantages of this option are its simplicity and initial credibility. It could be implemented with the stroke of a pen. It would signal at once the authorities' commitment to unification. On the minus side, however, it would not address at all the objective of mitigating the pain while efficiency gains materialize. The large devaluation associated with the instantaneous unification of the exchange rate at 24-to-one would have large up-front fiscal, redistributive, reallocation, and inflationary costs. The economic and political consequences could thus be so traumatic as to render the whole experiment unviable.

Option Two can be labeled *sector-by-sector gradualism*. It would consist in gradually depreciating the one-to-one exchange rate towards 24-to-one but on a sector-by-sector basis, that is, in different degrees and at different speeds for different sectors. On the plus side, by spreading the pain by sectors and over time, this option could be less traumatic than Option One. In addition, it would the authorities some scope for experimentation, hence better control over the entire unification process. On the negative side, however, this option would also spread the gains over time and unevenly across sectors. By delaying relative price changes across sectors (hence the supply response) it would limit the scope for efficiency gains. Thus, this option would meet the objective of mitigating the pain better than a *raw big bang* but would fall significantly short as regards the objective of front-loading and maximizing the efficiency gains. In addition, by giving the authorities a wide margin for discretion in the sector-by-sector implementation of the unification process, this option would be affected by credibility and uncertainty problems. It would in particular give rise to pressures to postpone the exchange rate unification for the most distorted sectors. At the extreme, therefore, this option might raise doubts as to whether the unification reform will ever be completed. Such uncertainties could promote a wait-and-see attitude, thereby further delaying the supply response. Last (but not least), the multiplicity of (sector-by-sector) exchange rates in the transition would further segment markets and distort price signals, impeding efficient resource allocation across sectors.

Option Three may be labeled *economy-wide gradualism*. It would consist in preannouncing a path of gradual convergence of the one-to-one rate to the 24-to-one rate, which would apply uniformly to all sectors, actors, and activities in the economy. By spreading the pain over time, this option would be less traumatic than a *raw big bang* and, because of its uniform, economy-wide application, it would be "cleaner" and less distortive

than a *sector-by-sector gradualism*. An *economy-wide gradualism* would do a better job than a *sector-by-sector gradualism* at frontloading and boosting efficiency gains. However, it would also give rise to policy uncertainty, which could lead to self-fulfilling failure. By spreading the adjustment over time, investors under this option could also adopt a wait-and-see attitude that would postpone the materialization of efficiency gains. In turn, the delay in the supply response would raise transition costs and may eventually induce the authorities to abandon the preannounced exchange rate unification path. This would end up validating the initial skepticism, turning it into a self-fulfilling prophecy.

Option Four may be labeled *fiscally-cushioned big bang*. Under this option, the two exchange rates would be unified on day one at the (new) rate of 24 pesos per dollar. However, to mitigate the initial pain, the shadow taxes and subsidies implicit in the dual exchange rate regime would be replaced by explicit, revenue neutral, non-distortionary *lump-sum* taxes and subsidies. The latter would apply to each enterprise that is currently conducting transactions at the one-to-one rate, and would be set at an amount that would fully neutralize the initial impact of the exchange rate unification on the income of such enterprises.⁷

To better illustrate the effect of the fiscal cushion, consider the example of an enterprise that sold Y dollars to the central bank at the one-to-one rate during a full calendar year prior to the unification (hence received Y pesos in exchange). After the unification, and in the absence of compensating lump-sum taxes, that enterprise would receive $24xY$ pesos, a major windfall in income. Hence, to avoid such an extraordinary jump in its revenues, during the first year immediately after the unification, the enterprise would be subject to a lump-sum tax of $23xY$ pesos, leaving it with the same Y pesos in revenue after the unification as it would have had before unification. The lump-sum tax would then be gradually phased out over a period of n years in line with a preannounced schedule. The tax might decline, say, to $15xY$ pesos on year two, $7xY$ pesos in year three, and might be altogether eliminated in year four. Similarly, an enterprise that benefitted from purchases of Z dollars at the one-to-one rate during a full calendar year prior to unification would receive a lump-sum subsidy equal to $23xZ$ pesos during a full calendar year after unification. Such a subsidy would ensure that the revenue of the enterprise is unaffected by the unification in the first year. As in the case of the lump-sum taxes, the lump-sum subsidies would subsequently be gradually phased out in line with the preannounced timetable.

Compared to the previous options, the option of a *fiscally-cushioned big bang* would be superior at maximizing efficiency gains because it would allow from the start all economic actors (both old and new) to operate under a new set of relative prices, hence under market-oriented and efficiency-enhancing incentives. Indeed, provided that appropriate safeguards are put in place (more on this below), the enterprises' decisions on investment and production would be totally unaffected by the lump sum taxes and subsidies. The amount of such taxes and subsidies would be set in absolute terms and

⁷ The amount of, say, the annual subsidy or tax implicit in the exchange rate spread that is received or paid, respectively, by each individual enterprise is known because the central bank keeps records on the amounts of CUCs or dollars sold to or purchased from each enterprise.

would not vary with the intensity of the profit-seeking production and investment effort of the enterprises involved. Moreover, a fiscally-cushioned big bang would be far more effective than the previous three options in shielding the economy from the transition pains, first by neutralizing them through lump-sum taxes and subsidies and then by offsetting (much if not all of) their impact through efficiency gains. Thus, this option would much better address the pain/gain balance. It would require, however, an adequate and non-trivial preparation and concomitant changes in economic policy, as discussed in Section 7. Before doing so, however, in the next two sections we illustrate the basic features of a *fiscally-cushioned big bang* by sketching how it would be implemented in the case two key sectors of the Cuban economy: the foreign-managed tourism industry and the sector of importing state enterprises.

5. The foreign-managed tourism industry

The current system for foreign-owned (or foreign-managed) tourism services entails very large efficiency losses. Because foreign hotel operators must pay for labor in dollars (or in CUCs, that is, at a one-to-one exchange rate) but the employee receives her payment in CUCs (i.e., at the 24-to-one rate), the spread between the two rates implies a heavy tax on local labor. To make this scheme feasible, the foreign hotel operator makes labor payments not directly to the worker but to a state-managed employment agency. Out of every dollar paid by the hotel operator to the employment agency, the worker receives only 1/24 dollars (about 4 cents), with the state retaining the remaining 23/24 dollars (about 96 cents) as a tax. While this arguably allows the Cuban government to capture and redistribute much of the rents generated by the tourism industry, the implicit tax distorts the allocation of labor in a major way and the high labor costs faced by the hotel operator severely discourages job creation, undermines the quality of hotel services, hinders new FDI, reduces tourism inflows, and promotes stealth employment (direct but not legal sale of labor services) in the tourism sector.

A *raw big bang* unification would be traumatic on impact because of its major redistributive effects. As long as average wages remain where they are today (which is likely to be the case if hidden unemployment remains prevalent in Cuba) and hotel prices do not decline (which is likely to be the case as long as competitive pressures do not build up), foreign operators would effectively benefit from rents conducive to large windfall profits. Conversely, fiscal revenues would collapse, thereby undermining the government's ability to compensate the losers and generating (via a widening fiscal deficit) demand-side upward pressures on prices. Foreign operators would have more dollars to repatriate and this would weaken the balance of payments (BOP) to the extent that such repatriation is not compensated by new FDI inflows, which would take time to materialize.

The windfall gains that would accrue to the foreign-managed tourism industry under a *raw big bang* could be initially and fully neutralized under a *fiscally-cushioned big bang* through a dollar-equivalent lump-sum tax paid by each hotel operator. The amount of the tax would vary from hotel to hotel and would be equivalent to the pre-unification spread between the two exchange rates (24 – 1) times the value of the foreign exchange that each hotel sold to the central bank at the one-to-one rate during the year prior to

unification. Once set for the first year, the lump-sum taxes would be gradually phased out over a predetermined period according to a preannounced phase-down schedule. Because each operator would have to pay the tax no matter what and the amount to be paid would be independent of what the operator does (or does not do) in Cuba, the tax would be non-distortionary (i.e., it would have no impact on the operator's investment or production decisions). For the system to work, however, the operator should not have the option to default on its tax obligations (for example by selling out its Cuban venture and exiting the country). The simplest way to prevent such default is via internationally binding contracts, so that operators would have to honor their commitments whether they leave Cuba or not.

As long as competitive pressures remain moderate, the exchange rate unification, by lowering labor costs, should boost the *marginal profitability* of both existing investments (i.e., expanded hotel occupation) and new investments (i.e., expanded capacity). Thus, existing enterprises should expand capacity and improve quality of service so as to secure and expand their market share in anticipation of greater competition. On impact (i.e., given the current stock of FDI), employment and service quality should rise, thereby raising tourism inflows. As a result, BOP and fiscal pressures should ease. Over time, the new rules of the game should boost new FDI and promote local investment, further benefitting the BOP and fiscal accounts. This would in turn allow for a smooth phasing down of the lump-sum taxes. Over the longer term, labor productivity should steadily increase, allowing for real wages to rise and the real exchange rate to appreciate, as happened in the transition towards a market economy of the Eastern European countries.

While the existing hotel operators would not face any competitive disadvantage at the margin (both new and old would compete under the same plain level field in this sense), they could object to the deal by arguing that the lump-sum taxes would erode their *average profitability*. They could fear that as new entrants come in and/or existing operators start competing more aggressively against each other, hotel prices would fall, thereby eroding their average (after lump-sum tax) profit margins. Yet, there are at least three strong counter-arguments to this line of reasoning. First, because the existing hotel operators can take full advantage of their installed capacity and knowledge of the local market, they should be the main beneficiaries of the initial boost in profitability and can position themselves to take the best advantage of any new business opportunity that becomes available as a result of unification. Second, even if profit margins become compressed due to increased competition, volume increases should offset much, if not all, of the decline in profit margins. Third, even in the worst-case scenario of a transient decline in the average profitability of existing Cuban hotels (due to the payment of lump-sum taxes in a more competitive environment), because the current owners (or managers) of these hotels are mostly large international chains, they should be able to absorb this decline without much impact on their bottom line.

6. State-owned importing industries

Consider now the case of the state enterprises that import "basic" goods (machinery and equipment, inputs, or consumption goods that are considered essential to the welfare of the low-income population). Again, the current system entails huge efficiency losses.

Using the one-to-one exchange rate for “basic” imports and the 24-to-one for wages amounts to a tax on labor that subsidizes “basic” imports and penalizes the consumption of “non-basic” imports. While this may contribute to evening out consumption across Cubans, the implicit subsidization weakens the central government’s finances and penalizes the employment-generating local production of import substitutes, thereby pressuring the BOP. Exchange rate unification should thus result over time in a much improved allocation of resources as enterprises take full advantage of the country’s relatively less costly local labor force and other comparative advantages.

However, as in the case of the tourism industry, a *raw big bang* could be quite traumatic on impact, especially if the state enterprises respond to the unification as private enterprises would, i.e., by seeking to cut their losses. The enterprises would raise prices, thereby triggering supply-side inflation, and cut down on employment in the face of declining demand, thereby exerting contractionary pressures on the economy. While this would improve the BOP and the finances of the central government, it would raise unemployment and erode the purchasing power of low-paid workers.

Alternatively (and perhaps more realistically), the initial impact under a *raw big bang* might only affect the public enterprises’ accounts and nothing else if these enterprises choose to respond passively. The enterprises in this case would continue importing and producing in the same fashion as before the unification but just run large deficits that would need to be covered by high and now explicit subsidies (as the subsidies implicit in the previous dual exchange rate system would come to the surface). The basic problem with this alternative scenario, however, is that granting explicit subsidies automatically would fundamentally undermine the central government finances (as the government would no longer receive the income from the exchange rate spread) while taking away the incentives for state enterprises to become more responsive to market signals, cut their costs, and improve their efficiency (business for them would basically continue as usual).

A *fiscally-cushioned big bang* unification would obviate the above problems by replacing the inefficient import subsidy (which is currently channeled via the dual exchange rate) with a dollar-equivalent lump-sum subsidy that initially fully neutralizes the fiscal, BOP, and inflationary impacts of the exchange rate unification. The lump-sum subsidy would vary from enterprise to enterprise and it would be set for the first (post-unification) year at an amount equal to the spread between the pre-unification exchange rates (24 – 1) times the total value of foreign exchange that each enterprise purchased from the central bank at the one-to-one rate during the year prior to unification. The lump-sum subsidies would be gradually phased out over a period of years in line with a preannounced schedule.

The gradual dismantling of the lump-sum subsidies would give state enterprises time to accommodate and would immediately enhance budgetary and public sector transparency. In particular, loss-making enterprises will come out fully into the open. The explicit subsidies should thus help promote accountability and generate incentives for state enterprises to restructure and improve efficiency in response to the changes in market conditions. For all of this to work well, of course, the government would have to resist the

pressures to extend and re-extend the timetable for the phasing out of the lump-sum subsidies. The fact that the new, unified exchange rate would apply to all foreign exchange transactions should help enhance competition, hence market discipline and responsiveness. All enterprises (existing state enterprises as well as potential new entrants) would operate from the start under the new rules of the game.⁸

For public enterprises to alter their behavior it would be crucial that the unification be accompanied, as soon as feasible, by appropriate reforms of their governance as well as by an in-depth reform of the scheme for subsidizing basic goods. Otherwise, the enterprises could continue conducting their business as usual. They could simply wait for the government to extend the subsidies indefinitely into the future under the threat of social upheaval. To prevent such an outcome, the state enterprises could be initially required, as a counterpart to the lump-sum subsidies they will receive, to set their pricing decisions so as to limit price increases, particularly on the most basic goods (indeed, the subsidies should make stable prices consistent with profitability). At the same time, to ensure that efficiency gains do materialize, the enterprises' production decisions should be nudged towards cost minimization and profit maximization. Even though efficiency gains would materialize as enterprises reorganize their operations under the new market prices, some price increases may be unavoidable as subsidies to enterprises are gradually phased out. A comprehensive reform of the public subsidy scheme will therefore also be needed in due course to protect the poor while limiting the fiscal costs (more on this below).

7. Habilitating reforms and sequencing issues

As noted, for the *fiscally-cushioned big bang* approach to work well, there would be a need for careful preparation, including the introduction (not necessarily before the unification but certainly as soon as feasible) of a number of crucial complementary reforms. The first area that would require reforms is the fiscal area. As noted, the shadow taxes and subsidies (implicit in the current dual exchange rate system) will first need to be replaced by lump-sum taxes and subsidies calculated for each enterprise affected by the exchange rate unification. Subsequently, however, the lump-sum taxes and subsidies would need to be themselves either eliminated and/or replaced by a well-designed and modern tax and subsidy system.

On the tax side, the government may need to revisit and adapt the tax system from the perspective of a new long-term, market-oriented environment. This may include first of all revising the taxes on FDI and foreign trade (import and export). In addition, to boost

⁸ There is also a hybrid option of staggering the *fiscally-cushioned big bang* unification over a period of time on a sector-by-sector basis. Thus, sector A (say, the tourism industry) could enter at once the single exchange rate regime while sector B (say, the state-owned importing industries or the exporting sector) could remain for a little while under the old regime. This option might give the authorities (or the affected enterprises in any given sector) more preparation time. It would be superior to Option Three (*economy-wide gradualism*) in that, even when applied on a sector-by-sector basis, the non-distortive nature of lump-sum taxes and subsidies would more effectively stimulate decisions to invest and raise production. As long as all enterprises in any given sector are brought in together into the unification scheme (instead of phasing in only a few, discretionally selected firms within a sector, while leaving other firms within the same sector shielded from competition), the resulting distortions might be limited and the market complications manageable.

revenue mobilization, a value added tax could be introduced while other internal taxes—especially the income tax on households and enterprises—are redesigned and modernized. *On the subsidy side*, the government may need to introduce a well-focused system of (cash or coupon) transfers targeted to the most basic goods and lowest income households. Unless this is done, the government could face an unsavory choice between a rapid but socially traumatic phasing-out of the lump-sum subsidy scheme and an ad-hoc prolongation of such, ultimately unsustainable, scheme.

In all cases, a *healthy fiscal position* would facilitate the unification of the exchange rate. A pre-unification fiscal surplus could allow the government to accumulate foreign reserves that could subsequently be used to “finance” temporary post-unification BOP deficits. A strong post-unification fiscal position would greatly ease monetary control and, hence, inflation.

At the same time, *public sector governance reforms* would be essential to ensure that state enterprises become more responsive to market signals. Enterprises should be given market-compatible mandates (i.e., cost minimization and profit maximization) and their performance assessed (and their managers rewarded) accordingly. Based on the newly emerging market prices, accounting practices should be simultaneously revisited to identify unviable state enterprises and facilitate their closure or restructuring. Inter-enterprise claims across balance sheets should be netted out and restructured as needed.

These governance reforms should ideally be accompanied (or followed as soon as possible) by *market-oriented reforms*. In particular, the entry of non-state firms should be encouraged to boost supply, both directly and indirectly (i.e., by exerting competitive pressures on state enterprises). It would also be desirable for efficiency purposes to disengage state firms from intermediating the business activities of private firms and citizens by, for example, allowing companies in the tourism sector that have foreign capital participation to select and hire their workers directly.⁹

To maintain inflation under control and allow the exchange rate to play a more substantial buffering role, the scheme will require the *gradual strengthening of the central bank’s management capacity and independence*. This in turn would require reforms to facilitate the gradual development of the interbank and exchange rate markets, as well as a sound and more vibrant financial intermediation.¹⁰

A *flexible exchange rate regime*, if feasible from the start, would in principle be desirable. In the short term, it would help identify the equilibrium post-unification exchange rate and reduce the risks of emerging exchange rate misalignments. Indeed, some appreciation of the exchange rate is likely to materialize as firms start to respond to the new exchange rate and market prices, which would lower the demand for imports and stimulate the creation of new exporting capacities. Over the longer term, further exchange

⁹ Under the fiscally-cushioned unification option, the fiscal loss that would result from allowing companies in the tourism sector to hire employees directly would be fully offset by the fiscal gain associated with the lump-sum tax on these companies.

¹⁰ See Vidal and Villanueva (2013).

rate appreciation pressures are likely to develop as efficiency gains translate into purchasing power gains. A flexible exchange rate system would thus have the benefit of allowing the real exchange rate to appreciate through nominal appreciation rather than inflation. At the same time, it would help absorb external shocks and limit de facto dollarization.

However, a flexible exchange rate regime may not be feasible in practice because the time required to establish the market and institutional infrastructure necessary for an independent monetary policy with exchange rate flexibility. Since waiting for these conditions to be in place would excessively delay the unification, it would make sense to envision a post-unification evolution towards exchange rate flexibility. Indeed, maintaining a fixed exchange rate immediately after unification might help stabilize expectations at a time of substantial, potentially discomfoting changes. In sum, it might thus be preferable to first unify the exchange rate regime and then move gradually over time to a more flexible exchange rate system. This process could follow a three-phased agenda of monetary reform:

- *Initial phase*: a fixed exchange rate with the central bank controlling credit expansion and excess liquidity in the interbank market.
- *Intermediate phase*: a somewhat flexible (but managed) exchange rate with monetary-aggregate targeting.
- *Final phase*: a monetary policy system based on inflation targeting and a flexible (yet still managed) exchange rate.

One final issue to be considered carefully when designing the exchange rate unification process is whether or not to also unify the currency. As noted earlier, currency unification is a conceptually and practically different issue than exchange rate unification. The issue of currency unification arises because Cuba has, in addition to two exchange rates, two currencies (the CUC and the CUP) that act as units of account, means of exchange, and stores of value.¹¹

When it comes to the issue of currency choice and unification, there are three main options: (i) formally adopting the dollar as the single currency; (ii) maintaining a dual currency system; and (iii) adopting the (new, post-unification) Cuban peso as the single currency. The pros and cons of these three options are as follows:

- *Full dollarization* would be unwise, as it would increase Cuba's vulnerability to adverse terms of trade or other large shocks, especially by magnifying their recessionary impact. Under such a scheme, it would rely on inflation or deflation as the main channels for real exchange rate adjustment. Given that shocks to Cuba are not necessarily symmetric to shocks to the United States (the country that issues of the dollar), the equilibrium real exchange rate in Cuba would need to adjust in directions that would not necessarily coincide with those needed by the U.S. economy. Hence, under full dollarization, Cuba

¹¹ Actually, Cuba has more than two currencies if the U.S. dollars or Euros that Cubans hold in cash are taken into account.

would lose the option to use its own currency as a policy instrument and shock absorber.

- *A dual currency system*, whereby the new (post-unification) peso coexists with the CUC may have some advantages. In this case, it would be better if CUC bills disappear and the CUC continues to exist solely as an electronic unit of account in which bank deposits and loans could be denominated. The CUC would retain the one-to-one relation and convertibility to the U.S. dollar, which would be in turn backed by a currency board arrangement. Maintaining the option of saving and lending in CUCs might provide incentives for persons and enterprises to use the CUC instead of the U.S. dollar or the Euro, thereby helping the economy re-monetize while confidence in the new Cuban peso builds up. However, as long as private citizens prefer U.S. dollars or Euros in cash, the shift of savings towards CUC-denominated bank deposits may be limited. At the same time, as long as state enterprises continue to be the main depositors, there is little risk of a depositor flight into dollar cash. Moreover, by fragmenting credits and deposits and reducing the depth of local-currency finance, the dual currency can hinder the strengthening of monetary management.
- In conclusion, *full peso-ization* is arguably the preferred option. Under this option, all CUCs would be mandatorily converted into new Cuban pesos at the outset, i.e., simultaneously with unification of the exchange rate.

REFERENCES

- Agenor, P.R. (1992), “Parallel Currency Markets in Developing Countries: Theory, Evidence, and Policy Implications”, Essays in International Finance No. 188, Princeton University.
- Agenor, P.R., and R. P. Flood (1992), “Unification of Foreign Exchange Markets”, IMF WP/92/32.
- Di Bella, G. and A. Wolfe (2008), “A Primer on Currency Unification and Exchange Rate Policy in Cuba: Lessons from Exchange Rate Unification in Transition Economies”, *Cuba in Transition—Volume 18*; Washington, Association for the Study of the Cuban Economy.
- Dreher, S. (2009), “Monetary Policy in Cuba”, University of Maryland.
- Feinberg, R. (2013), “Soft Landing in Cuba: Emerging Entrepreneurs and Middle Classes”, The Brookings Institution.
- Frenkel, J.A., and A. Razin (1986), “The Limited Viability of Dual Exchange Rate Regimes”, NBER WP No. 1902.
- IMF (International Monetary Fund). Various years. Annual Report on Exchange Rate Arrangements and Exchange Restrictions (AREAER). Washington, D.C.
- International Currency Analysis, Inc. Various issues. World Currency Yearbook. (formerly Pick’s Currency Yearbook). Brooklyn, N.Y.
- Kamin, S. (1991), “Argentina's Experience with Parallel Exchange Markets, 1981-1990” (No. 407). Board of Governors of the Federal Reserve System.
- Kiguel, M., and S.A O'Connell (1995), “Parallel exchange rates in developing countries”, *The World Bank Research Observer*, 10(1), 21-52.
- Lizondo, S. (1987), “Unification of Dual Exchange Markets”, *Journal of International Economics*, 22, pp. 57-77.
- Marion, N. P (1994), “Dual exchange rates in Europe and Latin America”, *The World Bank Economic Review*, 8(2), 213-245.
- Orro, R. (2008), “The Cuban Dual Monetary System and Challenges Ahead.” *Cuba in Transition--Volume 18*. Washington: Association for the Study of the Cuban Economy.
- Rogoff, K., Mody, A., Oomes, N., Brooks, R., and A.M. Husain. (2003). “Evolution and Performance of Exchange Rate Regimes”, International Monetary Fund.

Vidal, P. (2012a), “Monetary Duality in Cuba: Stages and Future Perspectives” in *Cuban Economic and Social Development: Policy Reforms and Challenges in the 21st Century*, David Rockefeller Center for Latin American Studies, Harvard University, Cambridge.

Vidal, P. (2012b), “Monetary and Exchange Rate Reform in Cuba: Lessons from Vietnam”, Institute of Development Economics, Japan External Trade Organization.

Vidal, P. and O. E. Pérez Villanueva (2013), “Política Monetaria y Reforma Bancario-Financiera,” Universidad Javeriana de Cali y Universidad de La Habana.