SECURE AFRICA PARTNERS

WORKING PAPER SERIES

# Exchange Rate Unification in Ethiopia

*Why Does the Foreign Exchange Premium Persist?*

By Sisay Lelissa

**WORKING PAPER** · PUBLISHED MAY 2026

## ABOUT THIS PAPER

This working paper diagnoses why Ethiopia's parallel foreign exchange premium re-emerged after the July 2024 unification reform, and sets out short-, medium-, and long-term priorities for locking in a durable, unified exchange rate.

It is independent policy research, drawing on econometric evidence, comparative reform experience from Nigeria, Angola, and Argentina, and a structured review of domestic and international media. The analysis and views presented are those of the author alone.


**AUTHOR**
Sisay Lelissa

**PUBLISHED**
May 2026

**CONTACT**
research@secureafricapartners.com


## CONTENTS


00 Executive Summary ..... 3

01 Context and Motivation ..... 5

02 Analytical Framework ..... 10

03 Problem Diagnosis: Causes of the Post-Reform Premium ..... 14

04 Empirical and Diagnostic Approach ..... 21

05 Policy Considerations ..... 40

R References ..... 45

A Appendix ..... 47

## Executive Summary

In July 2024, Ethiopia undertook one of the most consequential macroeconomic reforms in its recent history by unifying its foreign exchange (FX) regime under the Homegrown Economic Reform (HGER) agenda. Long-standing controls were dismantled, surrender requirements relaxed, and bi-weekly FX auctions introduced to restore market-based price discovery.

The immediate impact was dramatic. A sharp adjustment of the official rate cleared a large backlog of unmet FX demand and collapsed the parallel-market premium then exceeding 100 percent within weeks. For the first time in years, official and parallel rates briefly converged, confirming that the reform corrected a major price distortion. By mid-2025, however, the parallel premium had re-emerged in the range of 15–30 percent. While far smaller than pre-reform levels, its persistence raised a central policy question: *why did the premium return after unification, and what does this imply for the durability of unification?*

This paper argues that the resurgence does not reflect a reversal of reform intent or a failure of unification itself. Rather, it reflects transitional frictions arising from unresolved structural constraints, credibility and expectation dynamics, and incomplete market development. In short, the reform corrected the *price* of foreign exchange, but the *quantities, institutions, and expectations* underpinning a unified market remain in transition.

Drawing on econometric evidence, comparative reform experience, expert interviews and media-based analysis, the persistence of the post-unification premium is best understood as the interaction of three reinforcing forces.

- **First**, **FX scarcity has eased but not disappeared**. Reserves recovered from less than one month of import cover in mid-2024 to roughly two months by late-2025, reducing pressure but leaving excess demand unresolved. Econometric results confirm that reserve adequacy significantly lowers the premium, while rapid monetary base growth (M0) remains the dominant driver of FX demand. Price liberalization alone cannot clear the market when FX supply remains constrained and liquidity excessive.
- **Second**, **residual controls and shallow markets sustain segmentation.** Despite unification, residual policy restrictions continue to segment the FX market. The theoretical logic is direct: the parallel market premium is not a structural equilibrium phenomenon, if all

administrative restrictions on foreign exchange access were removed and the rate allowed to float freely, market arbitrage would drive official and parallel rates to a single clearing rate (Pinto, 1989; Montiel, 1986).

The premium is therefore policy-induced, not intrinsic to the Ethiopian economy. In Ethiopia's case, the remaining restrictions are substantial: the 2.5 percent NBE FX commission raises the cost of formal access; the capital account remains among the most restricted in Africa, with approximately 70–80 percent of capital account transactions subject to administrative controls, forcing legitimate portfolio diversification into informal channels; and administrative hurdles persist around dividend repatriation and invisible payments. These frictions sustain artificial segmentation between the official and parallel markets, compounded by shallow market structure: FX intermediation remains concentrated, interbank trading is thin, and hedging instruments are absent, encouraging precautionary FX hoarding.

- **Third**, **expectations and credibility drive short-run dynamics**. Econometric results show strong inertia in the premium, indicating that the sharp post-unification compression reflected a temporary credibility shock. These concerns were amplified by high inflation and fiscal tightening (including intensified tax collection), contributing to reform fatigue and discouraging cooperation with formal channels. In this environment, FX flows and remittances are more likely to shift back to informal networks, allowing brokers to regain relevance.

In sum, exchange rate unification corrected a major price distortion, but gaps in foreign exchange availability and credibility remain, explaining the return of the parallel premium. Whether a unified exchange rate endures will depend on how these gaps are addressed in practice.

1

SECTION 1

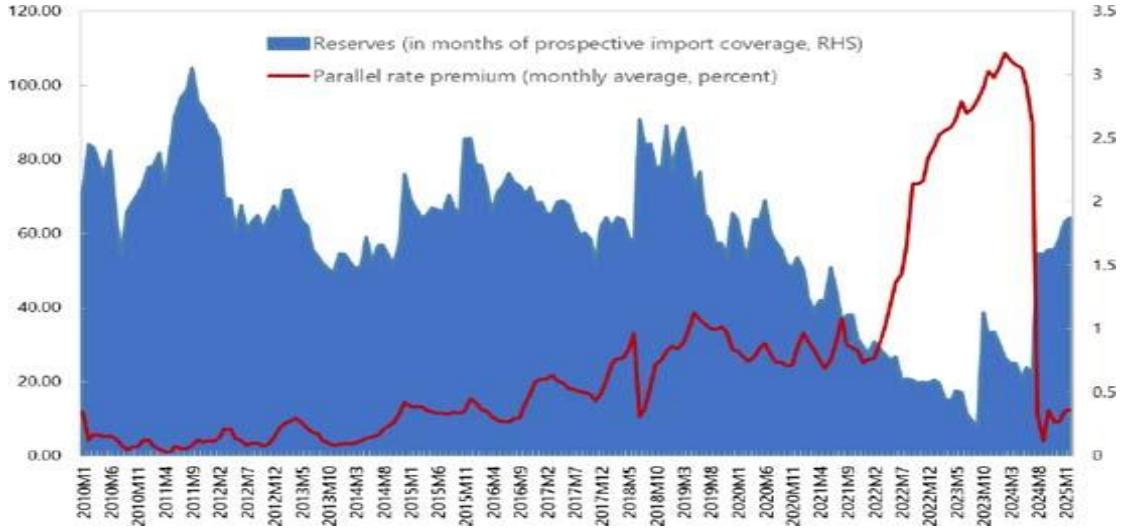
# Context and Motivation

Ethiopia’s persistent parallel FX premium has long been a barometer of macroeconomic stress. This section traces the pre-reform distortions and the July 2024 unification that sought to correct them.

## 1. Context and Motivation

Ethiopia's persistent parallel foreign exchange market premium (the gap between the official exchange rate and the parallel market rate) has long been a barometer of macroeconomic stress. By mid-2024, that premium had exceeded 100 percent, reflecting severe foreign currency shortages and an untenable overvaluation of the Birr. The widening gap signaled deep structural imbalances in Ethiopia's external sector, including chronic foreign exchange scarcity, administrative allocation of FX, and limited market mechanisms for price discovery.

**Figure 1: Reserve adequacy and Parallel Market Premium**




Source: IMF staff estimates, Gurhy et al. 2025)

An overvalued official rate not only distorted prices and hurt exports but also leads to rent-seeking and corruption in FX allocation (Pinto, 1989). In Ethiopia's case, access to foreign exchange was historically rationed through administrative mechanisms, creating a dual-market structure in which the parallel market became the primary channel for clearing excess demand.

Recognizing this, the Ethiopian authorities launched a major FX **unification reform on July 29, 2024**, as part of a broader Homegrown Economic Reform Agenda (HGER). The reform was intended to **eliminate the parallel market** and its harmful effects on the economy. Indeed, unifying the exchange rate was seen as a pillar of restoring macroeconomic stability, alongside debt restructuring , monetary policy reforms and fiscal adjustment.

Prior to the July 2024 reform, Ethiopia's foreign exchange system operated under a highly controlled institutional framework characterized by extensive administrative restrictions and limited market mechanisms. Although the exchange rate regime was officially classified as a managed float, in practice it functioned as a crawl-like arrangement in which the National Bank of Ethiopia (NBE) set the daily indicative exchange rate and tightly managed its adjustment. Ethiopia also remained under Article XIV transitional status under the IMF framework, allowing the continued use of exchange restrictions and multiple administrative controls on foreign exchange transactions. Foreign currency allocation was largely conducted through administrative rationing and priority import lists, while exporters were required to surrender the majority of their FX earnings to the central bank and commercial banks. At the same time, the capital account remained tightly restricted and the interbank FX market was shallow and underdeveloped. Together, these structural features created significant segmentation in the foreign exchange market and contributed to the emergence and persistence of a large parallel marketpremium prior to the reform.

**Table 1. Ethiopia's FX Regime Before and After the July 2024 Reform (IMF AREAER Framework)**

| Dimension | Institutional Features (Pre-Reform) |
|---|---|
| Exchange Rate Arrangement (IMF de facto) | Crawl-like arrangement characterized by gradual administrative depreciation of the Birr |
| Exchange Rate Regime (de jure) | Managed floating regime with the National Bank of Ethiopia (NBE) determining the daily indicative exchange rate |
| IMF Article Status | Article XIV (transitional arrangements allowing exchange restrictions and multiple currency practices) |
| Current Account Restrictions | Extensive foreign exchange rationing, priority import lists, and administrative FX allocation |

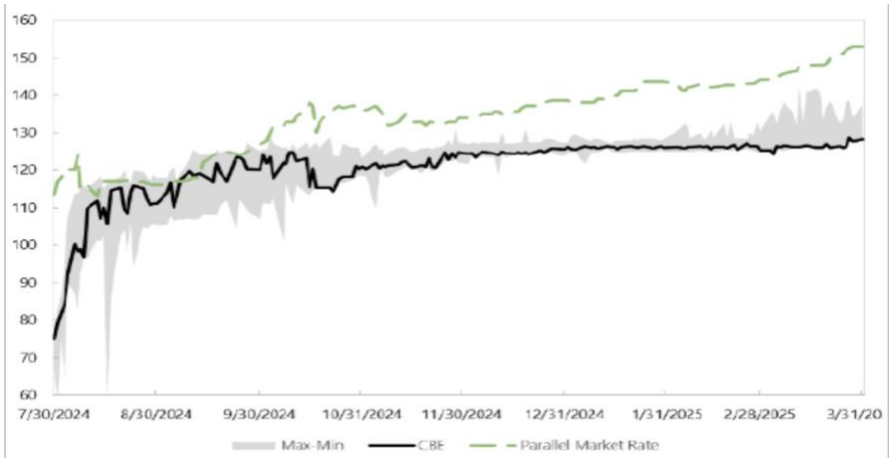
| Dimension | Institutional Features (Pre-Reform) |
|---|---|
| Capital Account Regime | Highly restricted capital account with strict controls on cross-border financial transactions |
| FX Allocation Mechanism | Administrative allocation through NBE priority sector lists and queue systems for importers |
| FX Surrender Requirements | Exporters required to surrender 50% of FX proceeds to NBE and 10% to commercial banks, retaining 40% in retention accounts |
| Foreign Exchange Market Structure | Limited interbank FX trading; most foreign exchange transactions intermediated through the central bank |
| Exchange Commission / Transaction Cost | 2.5% commission charged on foreign exchange sales by authorized banks |
| Monetary Policy Framework | Monetary aggregate targeting with reserve money as the operational target |

**Source:** International Monetary Fund (IMF), *AREAER* 2024

## Initial Progress and Renewed Concerns

The immediate impact of the July 2024 FX liberalization was dramatic. The National Bank of Ethiopia (NBE) allowed a one-off devaluation and held a special foreign exchange **auction** to discover a new market-clearing rate, moving the official rate from ETB 54/US$ to about **ETB 107.9/US$ in the first auction**. In subsequent bi-weekly auctions, the Birr was allowed to depreciate further in an orderly fashion; by June 2025 the official rate reached ETB 135/US$. In tandem, **surrender requirements were removed** whereby exporters were no longer forced to sell the bulk of FX to NBE, but could retain 50 percent indefinitely and sell the remainder at market rates. Import restrictions were lifted, ending the rationing that had hurt businesses. For a brief period, these bold measures nearly **closed the gap** by September 2024 the parallel rate converged close to the official rate (Figure 1). Policymakers and markets breathed a sigh of relief as the dual-rate system seemed to be ending.

**Figure 2: Range of Bank's Daily Weighted Average Rate (ETB/USD)**




Source: IMF Country Report, July 2025.

However, within months, cracks reappeared. By early 2025, reports showed the **parallel market premium creeping back into double digits**, reaching ~15 percent by April 2025. By July 2025, some market reports even indicated premiums of 25–30 percent on certain days (Figure 2). This resurgence was alarming: it implied that despite unification, segmentation of the FX market continued. The persistence of a parallel market, despite smaller than before, raised urgent questions: *What underlying factors are sustaining the premium? Why did the market not fully converge even after comprehensive reforms?* Answering these questions is crucial for Ethiopian authorities and their partners, as a lingering parallel market can undermine confidence in the reform program and stall economic recovery.

## Policy Significance

A sustained parallel premium matters for several reasons. **First**, it signals that macroeconomic imbalances are unresolved and typically, either a shortage of FX (external imbalance) or loss of confidence in the currency (internal imbalance), or both. If unaddressed, these pressures can fuel **inflation** (via expensive imports) and hamper growth (by deterring investment and diverting resources to unproductive arbitrage). **Second**, a parallel market undermines **policy credibility**: it indicates that many economic actors do not trust the official pricing of foreign exchange. In Ethiopia's case, this could erode the effectiveness of the new monetary policy framework, if people expect the Birr to keep depreciating beyond official forecasts, they may not respond to

interest rate signals as intended. **Third**, the premium has **distributional effects**. It tends to encourage illicit trades and "winners" who can access official FX at the low rate versus "losers" who must pay the parallel-market price, breeding perceptions of unfairness. This can weaken political support for reforms.

In sum, eliminating the parallel premium is not just a technocratic goal but a development imperative for Ethiopia's economy. The **timing** of this analysis is opportune. Eighteen months into the FX reform, the NBE and Ministry of Finance (MoF) need to decide on **next steps** to lock in the gains of unification and address the remaining gaps. This report aims to inform those next steps, providing a clear diagnosis of the problem and a feasible roadmap of solutions.

## 2. Analytical Framework

To ground the analysis, this section reviews two complementary strands of literature. The first examines theoretical and empirical research on parallel exchange rate premia, their causes, and macroeconomic impacts, including insights from currency crisis models and the concept of "fear of floating." The second surveys policy-oriented and case-study literature on exchange rate unification, focusing on sequencing debates (shock therapy versus gradualism), supporting policies, and institutional preconditions for success. Together, this literature provides a diagnostic framework and cautionary lessons directly relevant to Ethiopia's post-unification experience.

### 2.1 Theoretical Perspectives on FX Premia and Macroeconomic Impacts

#### Parallel Market Premia and Economic Distortions

A parallel market FX premium arises when the official exchange rate is maintained below the market-clearing level, typically due to exchange controls, capital restrictions, or pegged regimes that generate excess demand for foreign currency at the official rate. The result is a dual-price system: an official rate and a (higher) unofficial rate, with the premium commonly measured as the difference between parallel and official rate as a percentage of official rate.

The classic literature (e.g., Dornbusch *et al.*, 1983; Pinto, 1989) views a large premium as evidence of exchange rate misalignment and characterizes it as a distortionary system of implicit taxes and subsidies. Exporters are effectively taxed through surrender requirements at an undervalued

rate, while those with privileged access to official FX earn rents. Over time, these distortions shrink the formal traded sector: exports underperform, import-competing industries fail to develop (as imports are artificially cheap at the peg), and smuggling increases.

Easterly (2001) famously identified large parallel-market premia as among the "taxes on growth" in developing economies, alongside high inflation and trade barriers showing that they correlate with significantly weaker growth outcomes. A large premium also signals currency overvaluation and anticipated devaluation, deterring investment and encouraging capital flight or shifts into hard assets. Such premia typically coexist with broader macroeconomic instability, including fiscal imbalances and monetary instability, compounding uncertainty.

## First-Generation and Second-Generation Crisis Models.

Currency crisis models provide a useful framework for understanding persistent FX premia. In first-generation models (Krugman, 1979; Salant & Henderson, 1978), crises stem from an inconsistent policy mix and most notably monetary financing of fiscal deficits under a fixed exchange rate. Governments exhaust reserves defending the peg until a speculative attack forces collapse. In dual-rate systems, the parallel premium can be interpreted as anticipating this outcome: as deficits are monetized and the peg becomes unsustainable, the parallel rate depreciates sharply in expectation of eventual devaluation.

Second-generation models (e.g., Obstfeld, 1994; Flood & Marion, 1998) emphasize expectations, credibility, and multiple equilibria. Even when fundamentals are not decisively weak, doubts about policy sustainability that are often shaped by political economy considerations can trigger self-fulfilling crises. Applied to parallel markets, low credibility or fears of policy reversal can sustain a premium even absent immediate reserve exhaustion. Once established, parallel markets exhibit "expectational inertia": agents continue to reference the unofficial rate out of habit or caution.

## Fear of Floating and Policy Credibility

Calvo and Reinhart's (2002) concept of "fear of floating" is particularly relevant for emerging and developing economies. Many countries formally adopt floating regimes but intervene heavily to limit volatility, driven by concerns over inflation passthrough, balance-sheet effects in dollarized economies, or political costs of depreciation. This reluctance to fully allow price adjustment can

inadvertently sustain FX premia by slowing official rate adjustment or reintroducing minor controls.

The literature emphasizes that credibility is central. If market participants believe the central bank does not truly intend to let the exchange rate float, they may anticipate future policy reversals or further depreciation, reinforcing parallel market activity. In an IMF study of unification episodes, Gray (2021) argues that rapid, decisive unification is typically less costly than prolonged, cautious approaches and explicitly warning against fear-of-floating-driven gradualism. This insight is directly relevant to Ethiopia's post-reform strategy.

### Real Trade versus Portfolio Balance Models

Earlier theoretical work distinguishes between trade-based and portfolio-based explanations of parallel markets (e.g., Montiel, 1986; Agenor, 1992). In real trade models, the premium reflects *unmet import demand* and export diversion: parallel markets finance smuggled or unrecorded trade, with the premium equilibrating illicit flows. In portfolio-balance models, by contrast, FX demand reflects *store-of-value* motives. When domestic assets offer low or uncertain returns, residents hold foreign currency to hedge inflation and depreciation risk, embedding expectations and relative returns into the premium.

In high-inflation contexts, currency substitution becomes pronounced, with more stable foreign currency serving as a preferred savings vehicle. Portfolio models predict that interest rate differentials and expected depreciation strongly influence premia, an implication supported by empirical evidence showing that rapid monetary expansion and low real domestic returns are associated with larger and more persistent premia.

## 2.2 Policy and Empirical Literature on Exchange Rate Unification

A substantial policy literature examines how countries eliminate parallel markets and unify exchange rates. Core themes include the benefits of unification, debates over sequencing and speed, and the role of supporting macroeconomic and institutional policies.

### The Case for Unification

There is broad consensus that exchange rate unification yields significant welfare gains. Berg and Borensztein (2000) characterize dual exchange rate systems as rent-seeking, expensive, and

highly distortionary. Unification improves resource allocation by eliminating incentives for lobbying for FX access and increases fiscal transparency by removing quasi-fiscal subsidies embedded in overvalued official rates. Cross-country studies (e.g., Reinhart & Rogoff, 2004) show that countries that unified exchange rates and allowed greater flexibility achieved better inflation and growth outcomes than those that maintained prolonged dual regimes.

## Gradualism versus Shock ("Big Bang") Approaches

A central debate in the exchange rate literature focused on whether unification should occur through a decisive adjustment or a gradual approach. Crawling pegs or staged devaluations are often intended to limit inflation pass-through and allow time to rebuild reserves. However, recent evidence suggests it is not the first-best strategy for eliminating parallel market premia. Gray (2021) shows that partial adjustment tends to sustain expectations of further depreciation, encouraging FX hoarding and prolonging parallel market activity and what he characterizes as "death by a thousand cuts." As a result, gradualism is increasingly viewed as a second-best response to binding constraints rather than an optimal policy choice, and only effective when supported by strong macro anchors and a clearly communicated endpoint.

## Supporting Policies and Institutional Preconditions

Cross-country research shows that exchange rate unification is more likely to be sustained when it is embedded in a broader macroeconomic stabilization package rather than implemented in isolation (e.g., Berg and Borensztein, 2000; Reinhart and Rogoff, 2004). The literature consistently emphasizes the importance of tight monetary policy to anchor inflation expectations following unification as well as credible fiscal financing to avoid renewed monetary expansion and reserve buffers or external support to stabilize expectations during the transition (Gray, 2021).

This literature underscores a central lesson: exchange rate unification is necessary but alone is insufficient. Durable convergence requires addressing both the FX quantity constraint and the credibility and institutional foundations of the regime. These insights directly motivate the problem diagnosis, empirical analysis and policy roadmap developed in the subsequent sections.

3

SECTION 3

# Problem Diagnosis: Causes of the Post-Reform Premium

Building on the analytical framework, this section applies the literature on parallel exchange markets to diagnose why Ethiopia’s premium re-emerged after unification.

# 3. Problem Diagnosis: Causes of the Post-Reform Premium

Building on the analytical framework established in Section 2, this section applies the insights from the literature on parallel exchange rate premia and exchange rate unification to Ethiopia's experience. The theoretical and empirical literature highlights two broad mechanisms through which parallel market premia persist even after price liberalization: (i) **quantity-side constraints**, where foreign exchange supply remains insufficient relative to demand at the official rate; and (ii) **credibility and coordination failures**, where credibility gaps, expectations, information frictions, and market structure prevent convergence even when prices are formally unified.

This section therefore diagnoses the re-emergence of Ethiopia's parallel premium through these two complementary lenses.

## 3.1 FX Access and Structural Imbalances

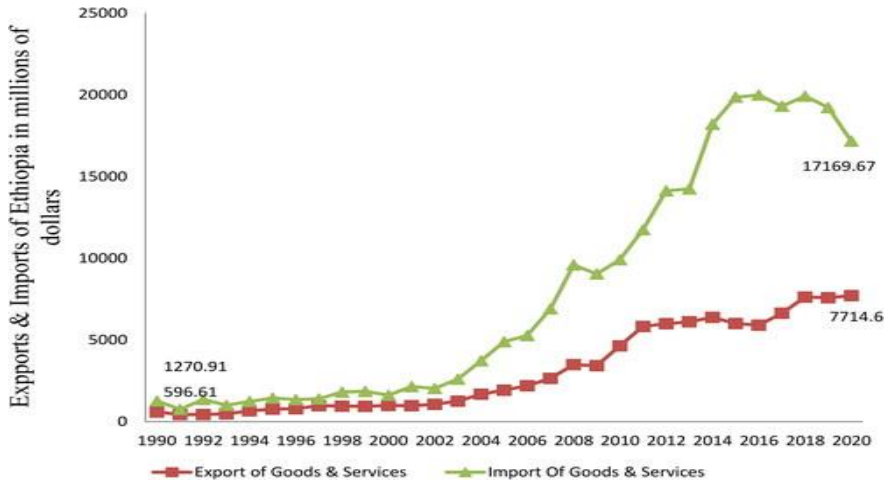
**Weak FX Inflows.** Ethiopia's foreign currency earnings remain structurally low relative to import needs. Exports of goods and services stood at about **9.6 percent of GDP** in FY2023/24, estimated to have risen to **12 percent in 2024/25**, but remain narrowly concentrated (coffee, gold) and vulnerable to shocks. **FDI inflows** have underperformed due to delayed privatizations and investor caution. Official aid declined sharply from around **12 percent of GDP in 2011 to under 4 percent in 2023**, reflecting donor retrenchment amid internal conflict. While **remittances are sizable**, a share continues to flow through informal channels when official conversion rules are perceived as unattractive.

**Figure 3: Ethiopia's Total Trade (US$ million)**



Source: (DOTS), 2024

**Figure 4: Ethiopia's Imports and Exports (US$ million)**

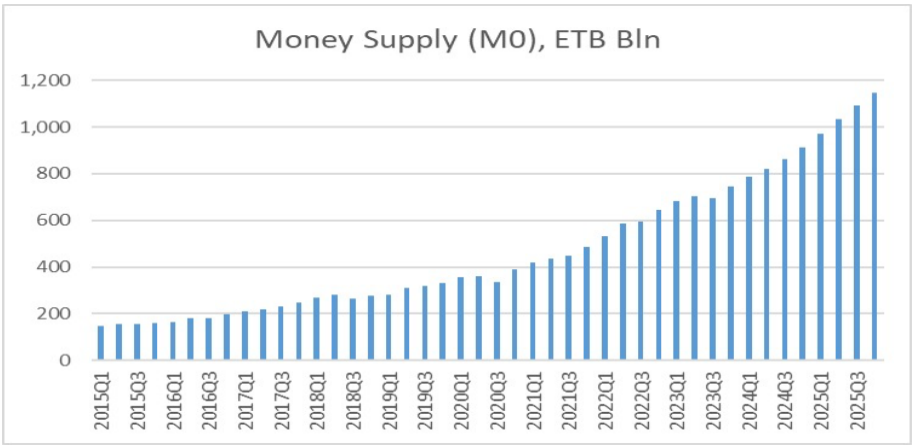



Source: Feyisa (2024), Cogent Economics & Finance

**Excess FX Demand and Import Pressure.** FX demand has remained robust due to development-related imports, population growth, and pent-up demand following the lifting of import bans in 2024. **Debt-service obligations** create episodic FX outflows, while precautionary and capital-flight motives add to demand. When official allocations remained constrained, unmet demand predictably spilled into the parallel market.

**Monetary Expansion and Excess Liquidity.** Prior fiscal monetization and directed credit led to rapid broad money growth over the past decade. Excess liquidity is best measured as the residual gap between money supply and money demand , a function of income growth and the cost of holding money (i.e., interest rates). When broad money grows faster than income and financial deepening can absorb, excess birr accumulates in the system and spills into FX demand. Although tighter monetary policy since late 2024 has slowed new injections, past excess liquidity continues to fuel FX demand.

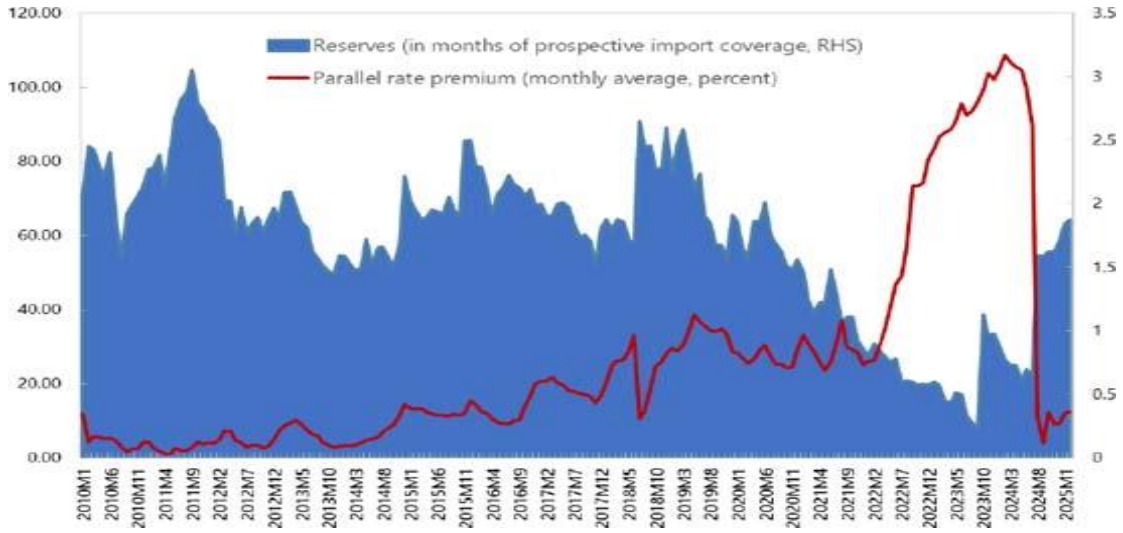
**Figure 5: Ethiopia's monetary base, ETB**




Source: National Bank of Ethiopia (NBE)

**Thin Reserve Buffer.** Gross reserves fell below **one month of imports** pre-reform, forcing strict rationing. Post-reform support lifted reserves to **US$3.8 billion by end-2024** and **US$5.15 billion by end-2025** (about **two months of import cover**), still below IMF's reserve adequacy metric of at least 3-4 months of imports of goods and services. Low reserves limit the authorities' ability to smooth shocks and amplify depreciation expectations, reinforcing parallel market pressures.

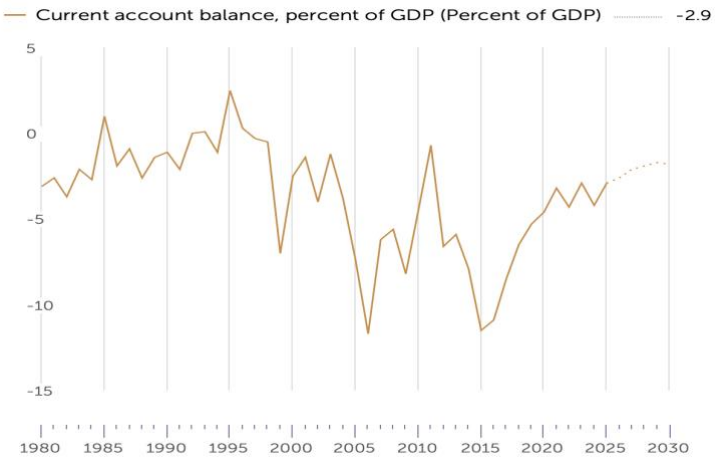
**Figure 6: Reserve adequacy and Parallel Market Premium**




Source: IMF staff estimates, Gurhy et al. 2025)

**Structural Current Account Gap.** Ethiopia's persistent current account deficit reflects a structural mismatch between import needs and FX earnings. Although the deficit narrowed modestly after the reform due to higher exports and earlier import compression, it remained in deficit in 2025. As long as the current account gap exceeds sustainable capital inflows, continuous pressure on the FX market and the parallel premium will persist.

**Figure 7: Current Account Balance**

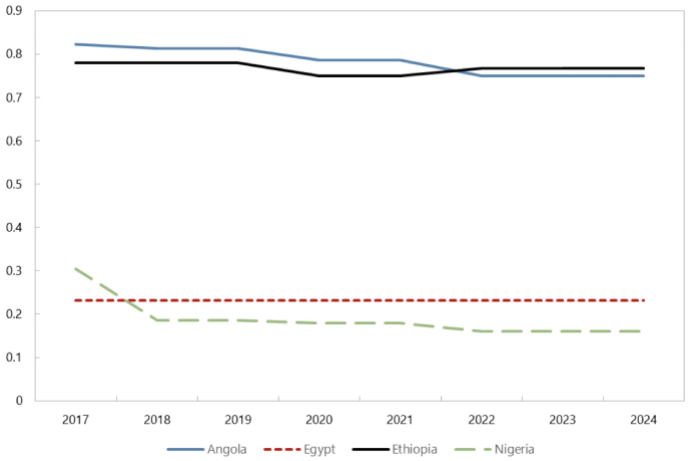



Source: World Bank , 2025

**Capital Controls and Pent-Up FX Demand.**

Ethiopia's stringent capital controls create pent-up demand for FX by preventing legal portfolio diversification and outward investment. This bottled-up demand often spills into informal channels, adding to parallel-market pressure. Managed, sequenced easing of capital controls could formalize part of this demand and help reduce the premium, rather than exacerbate it.

**Figure 8: Financial Account Restriction index**




Source: IMF staff estimates, Gurhy et al. 2025)

**Capital Controls and Pent-Up FX Demand.**

Ethiopia's stringent capital controls create pent-up demand for FX by preventing legal portfolio diversification and outward investment. This bottled-up demand often spills into informal channels, adding to parallel-market pressure. Managed, sequenced easing of capital controls could formalize part of this demand and help reduce the premium, rather than exacerbate it.

## 3.2 Credibility, Expectations, and Market Structure

***Policy Credibility and Consistency*.** Ethiopia's July 2024 unification initially boosted credibility in the new FX regime, as the NBE's commitment to market-determined rates represented a clear break from the managed peg. The July 2024 reform initially boosted credibility. Over time, however, mixed signals (such as, policy slippages relative to program targets, unpopular tax measures, etc.) have since weakened confidence, reviving doubts about the durability of the reform and the sustainability of the flexible rate.

***Depreciation Expectations.*** In Ethiopia's case, depreciation expectations have been self-reinforcing. Despite declining inflation through 2024, depreciation continued, and delays in disbursement of external financing reinforced market beliefs that further depreciation was likely. The re-emergence of a parallel premium itself fed these expectations, incentivising FX hoarding and perpetuating the very premium agents were anticipating.

***Information and Coordination Limits*.** The NBE has improved FX management transparency through regular bi-weekly auctions introduced in early 2025, but official rate information remains poorly visible beyond central bank and commercial bank channels. through regular auctions, but official rate information is still not widely visible beyond central bank channels. This limited outreach has reduced the NBE's ability to serve as a coordination focal point, allowing informal price signals including parallel market quotes circulated through social media — to continue shaping expectations and participation in Ethiopia's FX market.

***Market Structure Constraints*.** In Ethiopia, FX intermediation remains dominated by the Commercial Bank of Ethiopia, interbank trading is thin, and formal hedging instruments are entirely absent. Without formal risk-management tools, Ethiopian firms and households self-insure by accumulating FX balances, sustaining precautionary demand and reinforcing the role of the parallel market even as macro fundamentals improve.

**Figure 9: One Bank Concentration Ratio**

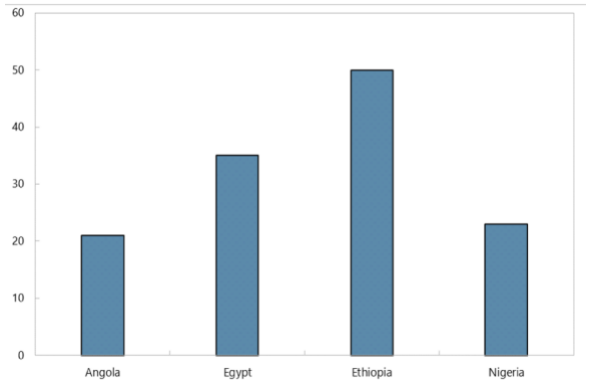
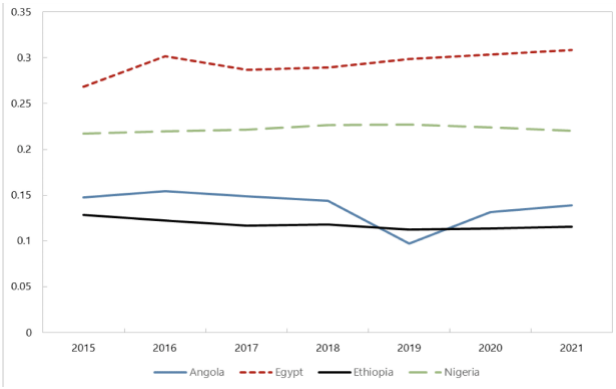



**Figure 10: Financial Development index**




Source: IMF staff estimates, Gurhy et al. 2025)

***Reform Fatigue, Political Economy, and Risk Perceptions.*** In Ethiopia, the July 2024 FX reform coincided with difficult fiscal adjustment under the IMF programme, contributing to reform fatigue, particularly among low-income households facing simultaneous import price increases and subsidy reductions. Skepticism toward externally supported programmes, rooted in Ethiopia's history of conditionality-linked reforms, further eroded domestic buy-in. Ongoing security concerns in Amhara and Oromia regions, and residual reputational damage from the Tigray conflict, added a country-risk premium to FX behavior, encouraging currency substitution and informal capital outflows that persist even as the macro reform agenda advances.

**Interplay of Quantity and Credibility Factors**

In Ethiopia, quantity and credibility failures have been mutually reinforcing. Low reserve buffers which stood at under two months of import cover through much of 2022–2024 weakened confidence in the NBE's ability to defend the new rate; excess broad money growth unanchored inflation expectations; and residual capital account restrictions simultaneously generated mistrust and pent-up FX demand. As a result, the persistence of the parallel premium in post-reform Ethiopia reflects a dense causal web rather than a single binding constraint. The subsequent regression analysis assesses the relative importance of these channels and identifies which have been most binding in Ethiopia's post-reform context.

## 4. Empirical and Diagnostic Approach

The analysis draws on multiple sources of evidence to diagnose why the parallel foreign exchange premium has persisted in Ethiopia and to identify practical policy implications. It combines three inputs: *econometric analysis of Ethiopia's FX dynamics*, *comparisons with recent exchange rate unification experiences in peer countries,* and a *structured review of domestic and international media*. Together, these sources are used to triangulate findings, assess credibility and market behavior, and link observed outcomes to concrete policy choices.

### 4.1 Comparative Case Studies

**Case Selection:**

To situate Ethiopia's post-unification experience within a broader policy context, this section reviews recent exchange rate unification and liberalization episodes in Angola, Nigeria, and Argentina. All three countries undertook major FX regime changes with IMF engagement and experienced differing degrees of success in eliminating parallel market premia. Their experiences provide concrete evidence on how policy credibility, reserve buffers, capital controls, and monetary conditions shape post-reform outcomes.

**Nigeria.** Nigeria attempted FX unification in 2016 and 2023 but failed to sustain convergence. Half-hearted liberalization, weak reserves (US$33bn in 2023), negative real interest rates, and inconsistent policy signaling allowed the parallel premium to re-emerge at around 20 percent

within months of each reform. The lesson: unification without credible monetary and fiscal commitment collapses.

**Angola.** Angola unified its exchange rate in 2018 but the parallel premium persisted at 20–30 percent for nearly three years before falling below 5 percent by mid-2021. Residual capital controls (covering ~80 percent of transactions), a concentrated banking system dominated by Sonangol, and delayed monetary tightening prolonged the adjustment. Premium convergence accelerated only after Angola removed remaining restrictions and raised real interest rates ,a sequencing lesson directly applicable to Ethiopia.

**Argentina.** Argentina unified twice in 2015 and 2025 with contrasting outcomes. The 2015 unification collapsed within three years because it lacked fiscal consolidation: deficit monetization eroded credibility and the premium re-emerged above 100 percent by 2023. The 2025 unification under President Milei succeeded because it was backed by a fiscal surplus, zero monetary financing, and US$20 billion in IMF support, compressing the premium to near-zero within days. The comparison shows that exchange rate unification is durable only when anchored by monetary and fiscal discipline.

Comparing all three cases, a pattern emerges: **the more comprehensive and credible the reform effort, the quicker the parallel premium fell** (Argentina showing both rapid initial success and subsequent relapse, Angola intermediate, Nigeria slow. The case studies thus reinforce the policy recommendations of this paper.

Detailed case narratives, country-specific evidence on premium dynamics, and policy sequencing analysis for all three countries are provided in Annex A10.

## 4.2 Econometric Evidence

### i. Long-Run and Short-Run Drivers of the Parallel Market Premium

The econometric analysis seeks to quantify the relationships between the parallel market premium and various economic drivers, in both the *short run and long run*. It distinguishes between structural, long-run drivers of the parallel market premium and short-run market behaviour/dynamics driven by shocks and expectations.

**Data and variables.** The analysis uses a quarterly dataset spanning 2015Q1–2025Q4, compiled from National Bank of Ethiopia (NBE) publications, Ministry of Finance data, and IMF International Financial Statistics. The parallel market premium is measured as the logarithm of the ratio of the parallel to the official exchange rate, drawn from the Reinhart–Rogoff Parallel Exchange Rate Database. Core macroeconomic variables include the monetary base (M0), foreign exchange reserves (months of import cover), inflation, the policy interest rate, and the current account balance (as a percent of GDP). A post-reform dummy (2024Q3) captures the FX unification regime shift. Interest rate and inflation differentials relative to the United States were also constructed as proxies for expectations and relative returns, though these showed weak explanatory power given closed capital account conditions. All series were tested for stationarity

## Long-Run (FMOLS) Model: Structural Determinants of the Premium

The long-run model is designed to identify the fundamental macroeconomic forces. Two FMOLS specifications are estimated: Model 2.a (M0 as monetary variable) and Model 2.b (inflation as an alternative proxy for excess liquidity, consistent with the monetarist approach in Grigorian, 2004). Model 2.a is the preferred specification given cleaner VIFs and theoretically expected coefficient signs that determine the level of the parallel market premium *over time*. Conceptually, it answers the question: *What level of premium should be expected given Ethiopia's underlying economic conditions?*

FMOLS is employed to address endogeneity and serial correlation inherent in cointegrating relationships, ensuring consistent long-run estimates. Two specifications are estimated: Model 2.a uses ln(M0) as the monetary liquidity proxy; Model 2.b replaces M0 with the domestic inflation rate, consistent with Grigorian (2004), who models excess liquidity as the residual between money supply and money demand, arguing that inflation serves as an effective proxy for this gap. Both models include the reserves-to-imports ratio, the current account balance (CA/GDP), and a post-reform dummy. The policy interest rate is excluded from both models due to severe multicollinearity with inflation (VIF > 25 when included). Together, the two models explain a substantial share of variation in the parallel market premium (Adj. $R^2$ = 0.779 for Model 2.a; 0.710 for Model 2.b).

*Model 2.a (M0 specification):* $\ln(\text{Premium}_t) = \beta_0 + \beta_1 \ln(M0_t) + \beta_2 (\text{Reserves/Imports}_t) + \beta_3 (\text{CA/GDP}_t) + \beta_4 D_{reform} + \varepsilon_t$

*Model 2.b (Inflation specification):* $\ln(\text{Premium}_t) = \beta_0 + \beta_1 \text{Inflation}_t + \beta_2 (\text{Reserves/Imports}_t) + \beta_3 (\text{CA/GDP}_t) + \beta_4 D_{reform} + \varepsilon_t$

where: $\text{Premium}_t$ = log ratio of parallel to official exchange rate; $M0_t$ = log of monetary base (ETB billions); $\text{Inflation}_t$ = domestic annual inflation rate; $\text{Reserves/Imports}_t$ = gross international reserves in months of import cover; $\text{CA/GDP}_t$ = current account balance as percent of GDP; $D_{reform}$ = dummy variable equal to 1 from 2024Q3 onward, capturing the FX unification regime shift.

**Table 2. Long-Run Cointegrating Equation (FMOLS)**

*Dependent variable: ln(Parallel Market Premium)*

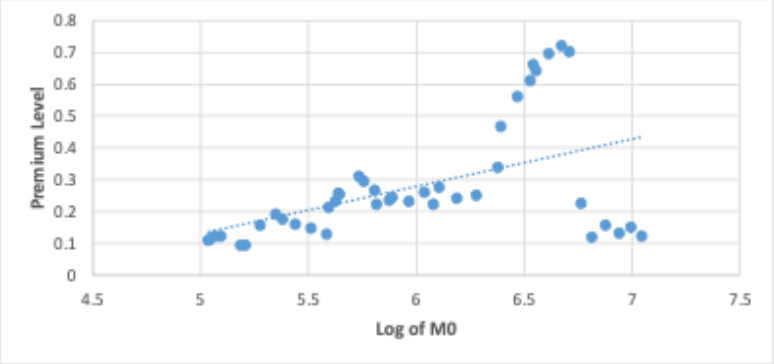
| Variable | Model 2.a (M0) | Model 2.b (Inflation) | Significance (2.a / 2.b) |
|---|---|---|---|
| ln(M0) | +0.35 | — | ***(p<0.01) / — |
| Inflation Rate | — | –1.11 | — / **(p<0.05) |
| Reserves/Imports (months) | –0.11 | –0.71 | n.s. / ***(p<0.01) |
| CA/GDP Ratio | –1.39 | +4.06 | n.s. / ***(p<0.01) |
| Post-2024Q2 Dummy | –0.49 | –0.19 | ***(p<0.01) / ***(p<0.01) |
| Constant | –1.77 | +1.24 | **(p<0.05) / ***(p<0.01) |
| *Observations = 43; Model 2.a: $R^2$ = 0.800, Adj. $R^2$ = 0.779; Model 2.b: $R^2$ = 0.738, Adj. $R^2$ = 0.710.* | | | |
| *Cointegration: Johansen* | | | |

| Variable | Model 2.a (M0) | Model 2.b (Inflation) | Significance (2.a / 2.b) |
|---|---|---|---|
| *trace test rank 1 at 5 percent level (both models). Newey–West correction applied. Residuals normally distributed. *** p<0.01; ** p<0.05; * p<0.10; n.s. = not significant.* | | | |

The long-run results point to a **monetary–liquidity model of the parallel premium**.

**Domestic liquidity is the dominant structural driver.** The positive and highly significant coefficient on the monetary base (M0) implies that sustained expansion of base money raises the equilibrium level of the parallel market premium. Given the scale of M0 growth over the past decade, this effect is economically large and confirms that excess birr liquidity consistently translates into FX pressure.

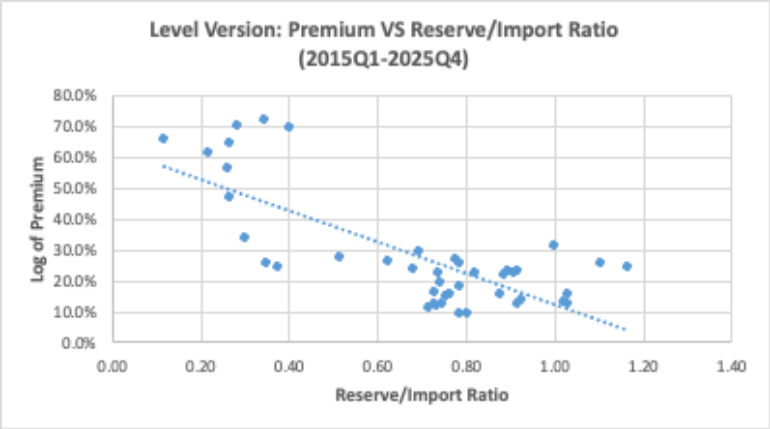
**Figure 11: Premium Vs domestic monetary base**




**Source: Own analysis**

**Reserve adequacy exerts a strong stabilizing effect.** Higher reserves relative to imports significantly reduce the long-run premium by easing foreign exchange scarcity and anchoring confidence. Ethiopia's post-reform reserve recovery therefore played a critical role in preventing a return to pre-reform premium levels.

**Figure 12: Premium Vs Reserves/import**




**Source: Own analysis**

**Inflation** is also a significant long-run driver in the alternative specification where it is used as a proxy for excess liquidity. The negative and statistically significant coefficient on the domestic inflation rate indicates that, in the long run, higher inflation is associated with a lower measured parallel market premium. This result should be interpreted carefully. It does not imply that inflation reduces foreign exchange pressures. Rather, it reflects the endogenous interaction between inflation and the official exchange rate in Ethiopia's post-reform environment.

Since July 2024, the official exchange rate has been allowed to adjust more flexibly, and part of the inflationary pressure has been absorbed through exchange rate depreciation. As the official rate depreciates in response to domestic price pressures, the gap between the official and parallel rates can narrow mechanically, compressing the measured premium even though underlying FX scarcity may persist. In this sense, a declining premium may partly reflect exchange rate adjustment rather than a genuine easing of foreign exchange constraints.

The **current account balance** is a significant long-run determinant in the inflation specification. The large positive coefficient indicates that external imbalances play an important role in sustaining the parallel market premium. Interpreted economically, a widening current account deficit corresponds with higher pressure in the foreign exchange market and therefore a larger parallel premium. This is intuitive: a persistent structural gap between import demand and foreign exchange earnings generates chronic excess demand for FX at the official rate, which spills into the parallel market. Ethiopia's structural current account deficit, driven by heavy

import dependence and a narrow export base concentrated in commodities such as coffee and gold, therefore acts as a fundamental long-run anchor of the premium that monetary and exchange rate policy alone cannot fully offset.

**FX unification shifted the equilibrium, but only partially.** The negative and significant reform dummy indicates that the July 2024 reform lowered the baseline premium by roughly 39 percent, even after controlling for fundamentals. This reflects improved price discovery and credibility. However, the magnitude of the effect also underscores that unification alone cannot eliminate the premium without sustained macroeconomic adjustment.

Overall, the long-run evidence shows that the premium persists because underlying liquidity and FX supply constraints persist. Structural reform reduced the level of the premium but did not eliminate its drivers.

## ii. Short-Run Dynamics and Adjustment (OLS)

**Model Specification**

To capture quarter-to-quarter movements in the premium during the transition period, two short-run OLS models are estimated in first differences: Model 2.a uses monetary base growth (M0) as the liquidity proxy, and Model 2.b replaces M0 with the domestic inflation rate, consistent with Grigorian (2004), who demonstrates that inflation can serve as an effective proxy for excess liquidity when monetary aggregates are subject to measurement or endogeneity concerns. This specification focuses on liquidity shocks, reserve changes, expectation-driven inertia, and temporary policy effects.

The short-run relationship is specified as:

$$\Delta \ln(\text{Premium}_t) = \alpha_0 + \alpha_1 \ln(\text{Premium}_{t-1}) + \alpha_2 \Delta(\text{Reserve/Imports}_t) + \alpha_3 \Delta M0_t + \alpha_4 \Delta \text{InflationRate}_t + \alpha_5 \Delta \text{PolicyRate}_t + \alpha_6 \Delta(\text{Current Account/GDP})_t + \alpha_7 D_{reform} + u_t$$

where: $\Delta \ln(\text{Premium}^t)$ = quarterly change in log parallel market premium; $\Delta \ln(\text{Premium}^t_{-1})$ = lagged premium growth, capturing inertia and momentum effects; $\Delta \ln(M0^t)$ = quarterly growth of the monetary base [Model 2.a]; $\Delta$ Inflation$^t$ = change in domestic inflation rate [Model 2.b]; $\Delta$(Reserves/Imports$^t$) = change in reserve adequacy; $\Delta$(CA/GDP$^t$) = change in current account balance as percent of GDP; $D_{-2024}Q_3$ = one-time dummy equal to 1 in 2024Q3, capturing the

immediate premium compression at unification; $u^t$ = error term. Note: this is an OLS model estimated in first differences, not an error correction model. The lagged dependent variable captures short-run momentum rather than adjustment toward a long-run equilibrium.

**Table 3. Short-Run Dynamics of the Parallel Market Premium (OLS)**

*Dependent variable: Δ ln(Premium)*

| Variable | Model 2.a (M0) | Model 2.b (Inflation) | Significance (2.a / 2.b) |
|---|---|---|---|
| Δ ln(Premium)(t−1) | +0.23 | +0.20 | **(p<0.05) / **(p<0.05) |
| Δ ln(M0) | +0.55 | — | **(p<0.05) / — |
| Δ Inflation Rate | — | +0.10 | — / n.s. |
| Δ Reserves/Imports | +0.11 | +0.05 | n.s. / n.s. |
| Δ CA/GDP | –0.55 | –0.35 | n.s. / n.s. |
| 2024Q3 Reform Dummy | –0.54 | –0.50 | ***(p<0.01) / ***(p<0.01) |
| Constant | –0.01 | +0.01 | n.s. / n.s. |

*Observations = 42; Model 2.a: $R^2$ = 0.767, Adj. $R^2$ = 0.735; Model 2.b: $R^2$ = 0.733, Adj. $R^2$ = 0.696. No heteroskedasticity (Breusch–Pagan); no serial correlation (Breusch–Godfrey); residuals normally distributed. CUSUM test confirms stability. *** p<0.01; ** p<0.05; * p<0.10; n.s. = not significant.*

**The short-run results highlight the transitional mechanics of the FX market.**

**Monetary conditions** are the dominant near-term driver. Short-run increases in monetary base growth transmit almost immediately into within-quarter increases in the parallel market premium, confirming that domestic liquidity is the primary pressure point at quarterly frequency. Inflation was tested as an alternative proxy for excess liquidity but proved insignificant in the short run, reinforcing that it is monetary base growth specifically, not inflation broadly that drives near-term FX pressure.

**Reserve adequacy** does not exert a statistically significant effect at quarterly frequency, a result that at first appears counterintuitive. The explanation lies in the nature of post-reform reserve accumulation: More fundamentally, reserve adequacy signals confidence and reduces structural FX scarcity over longer horizons and its effect does not manifest within a single quarter. The long-run estimates confirm that reserves matter, but through expectations and structural channels rather than within-quarter dynamics.

**The current account balance** similarly does not exert a significant short-run effect. The direction is as expected  where an improvement would ease FX pressure, but the effect is estimated at quarterly frequency, consistent with the well-established pattern that trade flows adjust slowly. The long-run results confirm that the structural current account deficit is a fundamental anchor of the equilibrium premium, but short-run stabilization must rely primarily on monetary and institutional measures rather than waiting for trade adjustment to materialize.

Two further results shape the short-run dynamics. The positive lagged premium coefficient indicates strong momentum effects: once the premium starts rising or falling, it tends to persist unless disrupted by a clear policy signal, reflecting the forward-looking nature of expectations in thin informal markets. The large negative reform dummy confirms that decisive policy action compresses the premium rapidly by shifting expectations. However, the presence of momentum implies that these credibility gains decay unless reinforced by consistent liquidity management and follow-through.

## Synthesis of Empirical Findings

Taken together, the long-run and short-run results provide a coherent explanation for the persistence of Ethiopia's parallel market premium.

• **Monetary expansion is the dominant driver across horizons.**

Monetary base growth (M0) emerges as the most powerful and consistent determinant of the premium in both models. In the long run, sustained expansion of the monetary base raises the equilibrium level of the parallel market premium. In the short run, increases in M0 are transmitted rapidly into FX market pressure within the same quarter. This suggests that domestic liquidity conditions are the primary near-term driver of exchange rate pressures. Consequently, strict control of monetary base growth becomes the most critical quantitative policy lever. Any renewed monetization of fiscal deficits would likely translate quickly into renewed FX market instability.

• **External current account imbalances anchor the premium structurally.**

A larger current account deficit is associated with a higher parallel premium, consistent with persistent FX scarcity in Ethiopia. Structural import dependence combined with a narrow export base generates chronic excess demand for foreign exchange, which spills over into the parallel market. Durable convergence of the premium therefore ultimately requires structural improvements in export capacity and a sustained narrowing of the external deficit.

• **FX reserves play a complementary stabilizing role.**

Higher reserve adequacy is associated with lower premiums in both the long-run and short-run estimations. In the long run, greater reserve buffers significantly reduce the premium, although the magnitude is not precisely estimated in all specifications. At the quarterly frequency, however, reserves are not statistically significant, suggesting that their influence operates primarily through confidence effects and longer-term fundamentals rather than through immediate market intervention.

• **Policy credibility and expectations influence adjustment dynamics.**

The reform dummy and the momentum term indicate that decisive policy actions can temporarily disrupt expectation-driven dynamics. This is evident in the sharp compression of the premium

following the July 2024 FX reform. However, the persistence captured by the momentum term suggests that credibility effects weaken over time unless reinforced by consistent policy follow-through. Delays in addressing remaining distortions risk allowing expectations to re-anchor around the parallel rate.

• **External adjustment operates slowly relative to financial variables.**

The limited short-run role of current account variables indicates that financial conditions such as liquidity, reserves, and policy signals that adjust much more rapidly than trade flows at quarterly frequency. As a result, short-term stabilization must rely primarily on financial and institutional policy measures, while export expansion and external rebalancing remain medium- to long-term solutions.

• **No single policy lever is sufficient.**

The evidence cautions against reliance on any single instrument. Durable convergence requires a coordinated policy mix: sustained monetary discipline, continued reserve rebuilding, removal of residual FX distortions, and clear policy communication to anchor expectations.

Overall, the econometric evidence reinforces the broader diagnostic conclusion: Ethiopia's parallel market premium is fundamentally a monetary and liquidity phenomenon, amplified by expectations and credibility dynamics. Locking in convergence therefore requires combining tight monetary control with reserve accumulation and credible, sequenced reforms.

## 4.3 Structured Media Review

**Rationale, Scope, and Method**

Understanding public narratives and market sentiment is essential alongside official data and policy announcements. Media sources shed light on how stakeholders interpret reforms, where information gaps or resistance arise, and how informal practices evolve. In Ethiopia's context, where expectations and perceptions strongly shape market behavior, media discourse provides valuable qualitative evidence on credibility, coordination, and compliance dynamics.

The review covers mid-2024 to early-2026 and draws on approximately 40–50 media items from at least 15 domestic and international outlets, including local newspapers (*Addis Fortune*, *The Reporter*, *Addis Standard*, *Birritu*), international media (Reuters, Bloomberg, CNBC Africa),

financial blogs and policy platforms, expert op-eds, and recorded interviews, speeches, and panel discussions on FX reform, inflation, and parallel market activity. IMF press releases and government statements are also included to assess official communication and signaling.

Media content is analyzed using a structured coding framework covering public sentiment toward reforms, perceived policymaker credibility, reported explanations for the parallel premium, observed market behaviors (such as hoarding or informal trading), and political reactions. The results are integrated throughout the diagnostic and discussion sections, particularly where expectations, coordination failures, and reform fatigue are most relevant, and they inform the assessment of policy feasibility by highlighting key political-economy and implementation risks.

This section distills themes from the media review and discusses their implications.

### I. Public Sentiment and "Reform Fatigue"

One clear theme in domestic media is a mix of hope and frustration among the public regarding the economic reforms. In mid-2024, when the FX reform launched, there was hopeful commentary: people saw previously unavailable goods appear in stores (since importers got forex at auction) and the parallel market rate fell, indicating progress. However, by 2025, sentiment shifted towards "reform fatigue." Ethiopians have endured high inflation (peaking over 30 percent), elimination of fuel subsidy, and aggressive tax collection.

Local outlets like *Addis Fortune* and *The Reporter* ran stories on the **cost-of-living crisis** persisting despite the reforms. An illustrative piece titled *"Beyond the Numbers: Ethiopia's Cost-of-Living Crisis Persists"* (Jan 2026) noted that while macro indicators improved, many citizens still struggled with expensive food and rent. This breeds skepticism: people ask, *"If the economy is 'improving' on paper, why is my daily life still hard?"* Such skepticism can erode trust in government announcements, including those about the exchange rate. In extreme cases, it can fuel conspiracy theories.

The notion of *IMF-backed reforms* often carries a stigma. Ethiopia's history (like many African countries) includes structural adjustment memories. So, the media sometimes framed new taxes or austerity as "IMF-imposed." For example, a piece in *The Africa Report* (2025) headlined "*Shrinking tax take puts IMF-backed reforms at risk* noted Ethiopia's tax-to-GDP was still low,

implicitly showing pressure for more taxes. Although directly unrelated issues with FX regime but in public eye it's one big program.

### II. Credibility of Institutions: National Bank and Government Communication

Media commentary frequently discusses the credibility of the NBE and the government's economic team. Historically, NBE was often perceived as operating with limited autonomy, particularly due to its role in financing government priorities. Under the Homegrown Economic Reform (HGER) agenda, this perception has begun to shift. Ongoing reforms, including the drafting of a new central bank law to strengthen institutional independence, especially through the appointment of a new board under the amended framework. While local media have paid limited attention to the legal reforms themselves, international outlets consistently frame them as important institutional milestones.

By contrast, domestic media coverage has focused more directly on NBE's operational responses to parallel market pressures. Reports such as Capital Ethiopia's coverage of "swift action against parallel forex market spikes" reflect a recognition that the central bank is actively engaging with market developments. While some commentary suggests that these actions are reactive in nature, the overall tone indicates growing respect for NBE's efforts to stabilize the FX market under difficult conditions. The narrative that emerges is of a central bank working to charter a new policy and deploy tools available at its disposal in a challenging transition period.

### III. Political backing and consistency

Media coverage consistently portrayed Ethiopia's government as a pro-reform actor, framing the July 2024 FX liberalization as a domestically led initiative under the Homegrown Economic Reform Agenda (HGER) rather than an externally imposed adjustment. This narrative of local ownership was widely reflected across domestic outlets, suggesting that the reform carried a degree of public legitimacy that IMF-backed programmes in other African contexts have historically lacked. However, media reporting also documented the limits of this framing. Political challenges including conflict in Amhara and Oromia, protests over cost-of-living pressures, and the lingering reputational effects of the Tigray war were consistently linked in coverage to weakened donor relations, reduced aid disbursements, and heightened investor uncertainty. Several outlets drew explicit connections between the security environment and reduced

appetite among foreign investors to commit capital. Taken together, the media record suggests that reform credibility in Ethiopia is not purely a function of macroeconomic policy settings but also conditioned by perceptions of political stability and governance, which shape whether economic agents view the FX regime as durable or reversible.

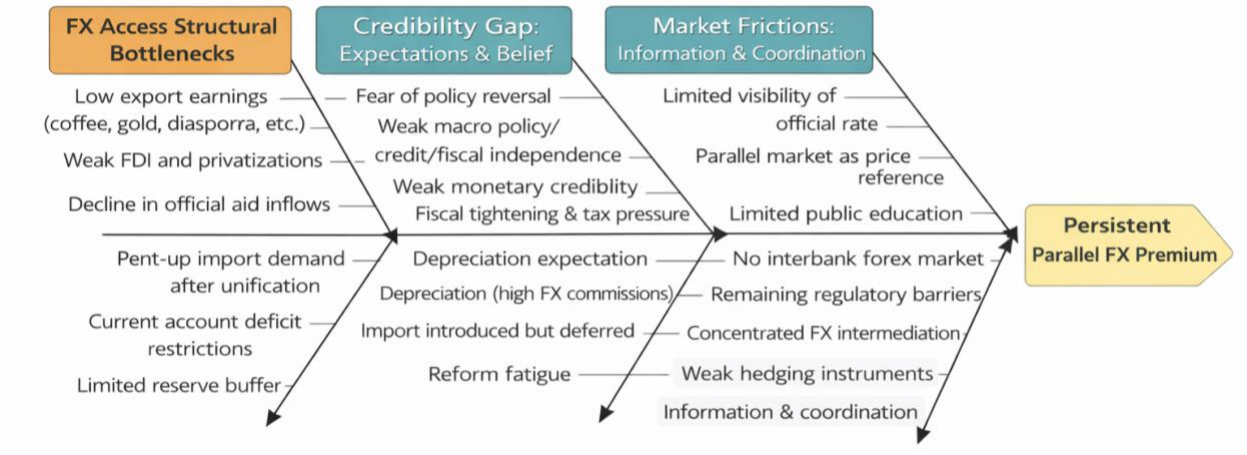
## 4.4 Integrated Analysis

This section consolidates the empirical and qualitative findings into an integrated diagnosis structured around the three main causal factors. Figure 15 presents a fishbone diagram illustrating how quantity constraints, credibility and expectations, and market-structure frictions interact to sustain the parallel market premium.

**Problem Statement:** Despite the July 2024 unification of exchange rates, Ethiopia's parallel-market FX premium persists (hovering around 20 percent in late 2025) instead of remaining closed. This indicates that the exchange rate reform, while initially effective, has not durably eliminated the gap between the official and curb market rates.

**Core Diagnosis:** The continued parallel premium is a *symptom* of three underlying, structural issues (1) FX access structural bottlenecks , (2) a Credibility deficit in policies and institutions, and (3) Information asymmetry and market inefficiencies. Each of these "bones" in the fishbone diagram contributes to the premium's resilience:

**Figure 13. Interaction of FX Access Structural Bottlenecks, Credibility Gaps, and Market Frictions in Sustaining the Parallel FX Premiu**




**Source:** Author's own analysis

1. **Structural Bottlenecks in FX Access and Policy Credibility Constraints:** A persistent imbalance exists between FX supply and demand at the official rate. However, a supply shortfall alone would not sustain a parallel premium under normal market conditions, it would simply depreciate the official rate. The premium persists because of structural bottlenecks in access to FX, administrative restrictions that prevent demand from clearing at the official rate, and policy credibility concerns that cause economic agents to seek the parallel channel. These factors are multi-faceted:
    - **Low FX Inflows:** Export revenues, though improving, remain too low to meet import needs. The economy is still heavily import-dependent (fuel, capital goods, etc.) while export capacity (coffee, gold, manufacturing) is limited and needs time to grow. Similarly, FDI and other capital inflows are only gradually recovering; the delay in debt restructuring also limited new inflows in 2024–25.
    - **Limited Reserves Buffer:** Gross international reserves, at ~2 months of import cover by end-2025, are still borderline. While better than <1 month in early 2024, this level is not enough to confidently meet surges in demand or cushion external shocks.
    - **Backlog of Unmet Demand:** Years of import restrictions and rationing created pent-up demand. After the unification, previously suppressed import needs flooded back. Even with the removal of bans on 38 categories, the economy could not immediately finance all desired imports through official channels without straining reserves. Some of this backlog spills into the parallel market (importers willing to pay a premium to get foreign goods now rather than queue).
    - **Ongoing Capital Controls:** While current-account transactions were liberalized, the capital account remains closed (e.g. residents cannot legally invest abroad, foreign investors face repatriation controls to an extent). This is prudent for stability, but it also means any residual desire to move capital (for diversification or fear of inflation) goes through unofficial channels, adding to parallel demand. Moreover, **net remittances** and informal flows can divert to the parallel market if senders or recipients find the official market cumbersome or rate unfavorable. Essentially, the formal market has not yet recaptured 100 percent of flows, some still leak to informal brokers.

- **Lagging Supply Responses:** Key supply-side reforms like attracting forex through diaspora bonds, accelerating privatizations (which could bring in FX), or increasing export volumes are in progress but take time. In the interim, the supply gap remains. For example, foreign banks entry and a new stock exchange (which could draw portfolio inflows) are part of HGER Phase II but not yet realized.

***Impact:*** As evidenced by the empirical analysis this quantity gap keeps upward pressure on the exchange rate. If the official market supply is constrained, not all demand is satisfied at the official price, and a portion is diverted to the parallel market at a higher price. The premium is essentially the market's way of "equilibrating" the excess demand that the official channel can't meet. Our long run model result of the premium's inverse correlation with reserves quantifies this: insufficient reserves (and by extension supply) lead to a higher premium. Thus, **as long as the FX shortage persists, a parallel premium will persist**.

2. **Credibility Deficit:** Economic agents lack full confidence in the stability of the new exchange rate regime and the supporting macroeconomic policies. This trust deficit manifests as:
    - **Fear of Policy Reversal:** After decades of tightly controlled FX policy, businesses and households find it hard to believe the reform is permanent. There is underlying anxiety that if the birr depreciates too much or if politics shifts, the government might reintroduce controls or a peg. This fear is reinforced by memories of past tight policies and by the knowledge that many countries with high inflation eventually re-impose controls. Even subtle signals can feed this.
    - **Weak Macro Policy Credibility:** High inflation (peaking ~33 percent in 2024) and large fiscal deficits in recent years have eroded trust in macroeconomic management. People ask: will the government truly tackle inflation and fiscal indiscipline? Until they see concrete, sustained action, they maintain expectations of birr depreciation. The **lagged premium inertia** is essentially an expectations phenomenon, if agents do not trust that inflation will be closed or that depreciation will halt, they extrapolate the trend.
    - **Institutional Uncertainty:** Ethiopia has undergone overlapping political and economic transitions in recent years. During a heightened phase of the reform process, the

resignation of the Governor of the NBE added to short-term uncertainty around policy continuity, even as the overall reform framework remained in place. In such environments, economic agents tend to hedge against perceived risks, which can temporarily amplify parallel market pressures through expectation-driven behavior.

- **Social and Political Pressures.** Public support for the reforms is fragile because the short-term costs have been highly visible. When the FX reform was launched in mid-2024, public sentiment was initially hopeful: the parallel rate fell and imported goods became more available. By 2025, however, this optimism gave way to reform fatigue. Households continued to face high living costs, fuel price increases, and heavier tax enforcement, even as official indicators improved. Domestic media increasingly highlighted this gap between macro progress and everyday hardship, raising skepticism about government claims that conditions were improving. These pressures weaken confidence in the permanence of the FX regime.
- **Market Participants' Skepticism:** Bankers, investors, and traders might also exhibit cautious behavior due to credibility concerns. For example, exporters might delay converting FX to birr expecting further depreciation (if they don't trust that inflation will be subdued or that current exchange rate is there to stay). Importers might over-order or stockpile goods (turning birr to inventory or dollars) anticipating future shortages or devaluation. Such behaviors, driven by lack of trust in stability, end up feeding the premium (more dollar demand now "just in case").

***Impact**:* The trust deficit amplifies the premium via **expectations and precautionary actions**. Essentially, even if fundamentals improve somewhat, a lack of credibility means people *behave* in ways that keep the premium up. Our model's finding that past premium influences future premium reflects this expectation spiral.

3. **Information Asymmetry and Market Frictions:** There is a gap in information dissemination and market infrastructure that has allowed the parallel market to act as a price leader and coordination tool:
    - **Lack of Transparent Price Data:** During the reform transition, real-time information on the official exchange rate was not widely accessible to the public. The NBE auctions

provided a rate, but it wasn't published daily in a way average businesses or individuals could easily observe. No unified tape of interbank trades existed yet. This vacuum meant that the parallel market's quotes broadcast through informal networks became the reference point. Even if the official rate and parallel weren't far apart at times, many people *only knew* the parallel quote (from, say, a popular Facebook page or messaging group). This one-sided information flow entrenches the parallel rate in expectations.

- **One-Way Expectations and Hording:** With incomplete info, people tend to rely on simple heuristics e.g. "the street rate is the real rate" or "last week dollar went up 5 birr, so it'll keep going". This herd behavior can sustain a premium beyond what fundamentals warrant. For instance, even when macro indicators improved in mid-2025 (lower inflation, higher exports), many still saw the parallel rate rising and assumed devaluation pressure was unabated, thus acting in ways (buying more FX) that kept the pressure on.
- **Market Structure Inefficiencies:** The FX market infrastructure is still developing. In late 2024–2025, reliance on twice-monthly auctions meant price adjustment was not continuous. If new information such as an increase in remittances or a debt relief deal came, the official rate might not instantly reflect it until the next auction. Such timing mismatches allow arbitrage and keep the parallel market relevant. Also, limited participation in auctions (banks only) and the nascent interbank trading meant the price might not fully reflect all demand-supply, whereas the parallel market, though illegal, brought in many players (individuals, unregulated dealers) giving an avenue for price discovery outside official channels. In short, the official market hasn't yet become *deep* and *liquid* enough to always set the clear price, especially early on, giving away that role to the parallel.
- **Informal Network Resilience:** Over years of rationing, an extensive informal broker network developed to intermediate FX. Those relationships and trust lines between brokers and traders, etc. don't vanish overnight. Many importers might still call their old broker first out of habit or convenience, even post-float, to compare rates or secure

funds quickly. If official channels don't strongly out-compete them in accessibility and reliability, these informal networks keep a foothold. Essentially, the **parallel market had a head-start in credibility** as the place to get FX easily, and it takes time for formal institutions to dislodge that.

- **Communication Gaps:** While the government and NBE did communicate about the reform, there may have been insufficient *public education* on what to expect from a unification. For example, if the public isn't informed that some volatility is normal and doesn't mean chaos, they might overreact to any rate movement by running to the parallel market. Also, if people aren't aware of new facilities (like FX bureaus) or new regulations (like exporters keeping 50 percent FX perhaps some thought old rules still in effect), they won't utilize them optimally. Any such communication gap can lead to suboptimal use of the formal market and persistence of old behaviors.

***Impact**:* Information and market asymmetries allow the parallel premium to *remain higher and more persistent than it would in a fully efficient market*. Thus, **lack of information flow and efficient markets created a wedge for the parallel market to survive**, where a more transparent system might have killed it off as soon as fundamentals moderately improved.

**Conclusion of Diagnosis:** Ethiopia's parallel market premium endures because of *structural problems* not simply because of minor policy errors or "speculators". It is essentially a **symptom** of structural macroeconomic imbalance, **exacerbated by** a **credibility gap** and **market frictions**. Therefore, any solution must attack all three roots: fill the FX supply gap (by allowing the FX rate to freely fluctuate to equate supply and demand) to remove the fundamental basis for a premium, build policy credibility to align expectations with fundamentals (so that people don't act in ways that create a premium), and improve market functioning and transparency so that any small premium is quickly arbitraged away and not allowed to become entrenched.

This integrated diagnosis directly shapes the policy recommendations in the next section, with each proposed measure targeted to one or more root causes to ensure a coordinated and effective response.

5

SECTION 5

# Policy Considerations

Short-term priorities to lock in credibility, medium-term reforms to deepen the market, and long-term anchors to prevent recurrence.

# 5 Policy Considerations

The fundamental benchmark for this policy agenda is clear: if all administrative restrictions on foreign exchange access were fully removed and the rate allowed to float without intervention, arbitrage would eliminate the parallel premium entirely (Pinto, 1989; Montiel, 1986). Each recommendation below is directed at closing the gap between Ethiopia's current regime and that benchmark.

The empirical results identify which policy levers matter and over which time horizon. In the long run, monetary base growth is the dominant driver of the equilibrium premium level, reserve adequacy exerts a significant stabilizing effect, and the structural current account deficit is a major long-run anchor of the premium. In the short run, domestic liquidity conditions transmit almost immediately into FX market pressure, while strong premium momentum indicates that credibility gains decay rapidly without consistent follow-through. The July 2024 reform confirmed that decisive policy action compresses the premium quickly by shifting market expectations but that compression is not self-sustaining. Reserve adequacy and current account conditions do not significantly affect the premium within a single quarter, confirming that these channels operate over longer structural horizons.

The policy response must therefore be sequenced to match these transmission channels: immediate actions targeting liquidity and credibility; medium-term reforms addressing market depth and capital controls; and long-term anchors ensuring monetary discipline and structural FX supply.

## A. Short-Term Priorities (Next 12–18 Months): Lock In Credibility

The short-run results identify monetary base growth and reform credibility as the two dominant fast-moving channels , both operate within a single quarter. Premium momentum further shows that credibility gains dissipate quickly, making consistent signaling as important as the initial reform. Residual administrative frictions act as coordination failures that keep the parallel market active even as fundamentals improve.

- **Maintain Tight Monetary Conditions.** The NBE should preserve a restrictive monetary stance until inflation is firmly in single digits and avoid any resumption of fiscal deficit monetization. The empirical results consistently identify domestic liquidity expansion as

the most powerful driver of the premium at both the short- and long-run horizons. Tight policy also reduces FX-inflation pass-through, allowing the real exchange rate to depreciate and gradually correcting the structural import-export imbalance that underlies persistent FX scarcity. Any premature relaxation would reignite both channels simultaneously. Feasibility: moderate but essential; politically sensitive.

- **Remove Current-Account Distortions.** The 2.5 percent NBE FX commission should be eliminated, limits on travel and invisible payments relaxed, and dividend and profit repatriation procedures streamlined. Residual transaction costs create artificial segmentation that prevents arbitrage from closing the official-parallel gap. The empirical results confirm that decisive, credible removal of distortions produces rapid and significant premium compression; conversely, leaving visible frictions in place signals incomplete liberalization and sustains the parallel market as a coordination point. Feasibility: high; regulatory adjustments only.
- **Strengthen Price Transparency.** Full auction outcomes including volumes, cut-off rates, and allocation rules should be published promptly, auction frequency increased, and the electronic interbank FX platform accelerated. The results show that premium dynamics are highly forward-looking and sensitive to credibility signals. Opaque or intermittent official price discovery allows informal market quotes to anchor expectations, sustaining the premium even when underlying fundamentals improve. Feasibility: high; technical groundwork underway.
- **Rule-Based FX Auction Smoothing.** The NBE should continue targeted auctions within available reserves to smooth temporary shortages and prevent disorderly conditions. While short-run reserve changes do not significantly affect the premium at quarterly frequency, the long-run results confirm that reserve adequacy exerts a meaningful stabilizing effect on the equilibrium premium. Rule-based, transparently framed auctions preserve this confidence channel without signaling a return to a fixed rate.

### B. Medium-Term Reforms (2–4 Years): Deepen the Market

The long-run results identify the structural current account deficit as a major anchor of the equilibrium premium, confirming that durable convergence ultimately requires improvements in structural FX supply, not just demand management. Reserve adequacy also matters significantly at the equilibrium level. These channels operate over multi-year horizons and require market-deepening reforms to address.

- **Sequenced Capital Account Easing.** Selected capital account restrictions which incudes travel allowances, and FDI exit rules should be incrementally relaxed once reserves and confidence strengthen. Rigid capital controls sustain the parallel market even after current-account liberalization by forcing legitimate diversification demand into informal channels, amplifying the structural FX scarcity that the long-run results identify. Sequenced easing reduces this pent-up demand while preserving macro stability.
- **Develop FX Hedging Instruments.** Basic FX forwards and swaps for trade-related transactions should be introduced and continuous interbank trading expanded. Precautionary FX hoarding by firms and households adds structural demand to the parallel market. However, hedging tools only function effectively once the rate is credibly market-determined and the empirical results confirm that credibility is the prerequisite for any market-deepening to work. Full rate credibility must precede hedging infrastructure.
- **Enhance FX Sector Competition.** FX bureau licensing should be expanded, fintech and money transfer operator participation encouraged, and foreign bank entry advanced under the revised banking law. Market concentration sustains a convenience premium for informal channels; greater competition narrows spreads, improves formal market access, and reduces the structural demand for parallel markets.
- **Rebuild Reserve Buffers.** The NBE should pursue IMF program targets for net international reserves and prioritize reserve accumulation in fiscal planning. The long-run results identify reserve adequacy as a significant structural determinant of the equilibrium premium, operating through confidence and FX availability channels.

Sustained accumulation is therefore a complement to, not a substitute for, monetary tightening.

### C. Long-Term Anchors (4+ Years): Prevent Recurrence

The long-run empirical results explain the majority of premium variation through structural monetary and external fundamentals, confirming that the equilibrium premium is not a short-run phenomenon. Preventing recurrence requires institutionalizing the conditions which entails monetary discipline, structural FX supply, and regime credibility, that the long-run model identifies as the equilibrium determinants.

- **Institutionalize Central Bank Independence.** The revised NBE law should be implemented, an independent board appointed, and interest-rate-based monetary policy entrenched. The long-run results confirm that monetary discipline and regime credibility are the two most important structural determinants of the equilibrium premium. Institutional independence removes the risk of discretionary monetary expansion and anchors the expectation that policy discipline will persist.
- **Build Structural FX Supply.** Export diversification, tourism recovery, and FDI should be accelerated through coordinated sectoral reforms. The long-run results identify the structural current account gap as the largest single determinant of the equilibrium premium. Short-term tools cannot substitute for durable export growth; only a sustained narrowing of the structural FX gap removes the fundamental demand pressure that sustains the parallel market.

## References


- Agénor, P.-R. (1992). Parallel currency markets in developing countries: Theory, evidence, and policy implications. *Princeton Studies in International Finance*, No. 73.
- Berg, A., & Borensztein, E. (2000). The choice of exchange rate regime and monetary policy in the aftermath of currency crises. *IMF Working Paper* No. 00/29.
- Calvo, G. A., & Reinhart, C. M. (2002). Fear of floating. *Quarterly Journal of Economics*, 117(2), 379–408.
- Dornbusch, R., Dantas, D., Pechman, C., Rocha, R., & Simões, D. (1983). The black market for dollars in Brazil. *Quarterly Journal of Economics*, 98(1), 25–40.
- Easterly, W. (2001). *The elusive quest for growth: Economists' adventures and misadventures in the tropics*. Cambridge, MA: MIT Press.
- Flood, R. P., & Marion, N. P. (1998). Perspectives on the recent currency crisis literature. *IMF Working Paper* No. 98/130.
- Gray, S. (2021). Exchange rate unification: Lessons from experience. *IMF Staff Discussion Note* No. 21/04.
- Grigorian, D. A., Khachatryan, A., & Sargsyan, G. (2004). Exchange rate, money, and wages: What is driving prices in Armenia? *IMF Working Paper No. WP/04/229*. International Monetary Fund.
- Gurhy, B., Lee, K., & Garimella, S. R. (2025). *Potential drivers of post-reform parallel market premium*. IMF Selected Issues Paper, No. 2025/105. Washington, DC: International Monetary Fund.
- Krugman, P. (1979). A model of balance-of-payments crises. *Journal of Money, Credit and Banking*, 11(3), 311–325.
- Montiel, P. J. (1986). Empirical analysis of the determinants of the parallel market premium. *World Bank Economic Review*, 1(1), 55–75.
- Pinto, B. (1989). Black market premia, exchange rate unification, and inflation in sub-Saharan Africa. *World Bank Economic Review*, 3(3), 321–338.

- Reinhart, C. M., & Rogoff, K. S. (2004). The modern history of exchange rate arrangements: A reinterpretation. *Quarterly Journal of Economics*, 119(1), 1–48.
- Salant, S. W., & Henderson, D. W. (1978). Market anticipations of government policies and the price of gold. *Journal of Political Economy*, The University of Chicago Press, 86(4), 627–664
- Tesfaye, E. (2025). Ethiopia's Exchange Rate Unification: The Double Inelasticity Trap, Banking Sector Arbitrage and Speculation, and the Reserve Adequacy Imperative (Working Paper). Unpublished manuscript.

# Appendix

*Parallel FX Market Premium in Ethiopia: Empirical Evidence (2015Q1–2025Q4)*

## Annex A1: Data Description and Variable Definitions

The dataset covers 44 quarterly observations from 2015Q1 to 2025Q4. All variables are sourced from public databases and national authorities.

| Variable | Code | Format | Source | Description |
|---|---|---|---|---|
| Parallel Market Premium | LN_PREMIUM | Log | Trading Economics | ln(Parallel Rate / Official Rate) |
| Monetary Base (M0) | LN_M0_SA | Log | NBE / Trading Economics | Log of seasonally adjusted monetary base (ETB) |
| Reserve Adequacy | RESERVE_IMPORTS | Ratio | Trading Economics | Gross international reserves divided by monthly imports |
| Current Account/GDP | CA_GDP_2 | Ratio | World Bank / NBE | Current account balance as percent of GDP |
| Domestic Inflation Rate | ETH_INFLATION | Percent | Trading Economics | Year-on-year CPI inflation rate (percent) |
| Policy Rate | ETH_POLICY_RATE | Percent | NBE | National Bank of Ethiopia monetary policy rate |
| Post-Reform Dummy | DUMMY_POST2024Q2 | Binary | Author | = 1 from 2024Q3 onward (FX unification reform) |
| Reform Quarter Dummy | D_2024Q3_SINGLE | Binary | Author | = 1 in 2024Q3 only (immediate reform shock) |

## Annex A2: Unit Root Tests (Stationarity)

All variables are tested for stationarity using the Augmented Dickey-Fuller (ADF) test with Schwarz information criterion (max lag = 4) and the Phillips-Perron (PP) test with Newey-West bandwidth. Tests include trend and intercept. All variables are integrated of order one, I(1), confirming the appropriateness of FMOLS for the long-run specification and first-differencing for the short-run OLS.

| Variable | ADF I(0) | ADF I(1) | PP I(0) | PP I(1) | Conclusion (5%) |
|---|---|---|---|---|---|
| Parallel Market Premium | 0.883 | 0.001 | 0.765 | 0.001 | I(1) |
| M0 (Monetary Base) | 0.400 | 0.000 | 0.373 | 0.000 | I(1) |
| CA/GDP Ratio | 0.131 | 0.000 | 0.195 | 0.000 | I(1) |
| Reserve/Imports Ratio | 0.590 | 0.000 | 0.545 | 0.000 | I(1) |
| Domestic Inflation Rate | 0.537 | 0.006 | 0.566 | 0.005 | I(1) |
| Policy Rate | 0.867 | 0.001 | 0.958 | 0.001 | I(1) |
| Net Domestic Assets | 0.623 | 0.009 | 0.669 | 0.001 | I(1) |
| Post-Reform Dummy | — | — | — | — | Deterministic |

*Note: p-values reported. $H_0$ = unit root. All series fail to reject $H_0$ at level and reject at first difference, confirming I(1). Period: 2015Q1–2025Q4.*

### Annex A3: Johansen Cointegration Tests

Johansen trace and maximum eigenvalue tests are conducted with lag intervals 1–4. Both statistics confirm the presence of at least one cointegrating relationship at the 5 percent significance level for the preferred models (Models 2.a and 2.b), validating the use of FMOLS for long-run estimation.

| Test | Model 2.a (M0) | Model 2.b (Inflation) | Model 3.a | Model 3.b |
|---|---|---|---|---|
| Trace Rank (5%) | 1 | 1 | 2 | 1 |
| Max-Eigenvalue Rank (5%) | 1 | 1 | 2 | 1 |
| Cointegration confirmed? | Yes | Yes | Yes | Yes |
| Lag interval | 1–4 | 1–4 | 1–4 | 1–4 |

*Note: $H_0$ = no cointegration. Rejection at 5% confirms long-run equilibrium relationship.*

### Annex A4: Long-Run Estimation Results (FMOLS)

The long-run models are estimated using Fully Modified Ordinary Least Squares (FMOLS) with Bartlett kernel and Newey-West fixed bandwidth (= 4.0). The dependent variable is the log parallel market premium [ln(Parallel Rate / Official Rate)]. Sample: 2015Q2–2025Q4 (43 observations). Cointegrating equation deterministics: constant and post-2024Q2 dummy.

**Table A4.1: Long-Run FMOLS Preferred Specifications (Models 2.a and 2.b)**

| Variable | Model 2.a Coeff. | Sig. | Variable Definition | Model 2.b Coeff. | Sig. | Variable Definition |
|---|---|---|---|---|---|---|
| ln(M0) | +0.35 | *** | Log monetary base | — | — | — |
| Inflation Rate | — | — | — | −1.11 | ** | Domestic CPI inflation |
| Reserves/Imports | −0.11 | n.s. | Reserve adequacy ratio | −0.71 | *** | Reserve adequacy ratio |
| CA/GDP Ratio | −1.39 | n.s. | Current account/GDP | +4.06 | *** | Current account/GDP |
| Post-2024Q2 Dummy | −0.49 | *** | FX reform regime shift | −0.19 | *** | FX reform regime shift |
| Constant | −1.77 | ** | — | +1.24 | *** | — |
| $R^2$ | 0.800 | | | 0.738 | | |
| Adjusted $R^2$ | 0.779 | | | 0.710 | | |
| Normality (Jarque-Bera) | Normal | | | Normal | | |
| Cointegration (Johansen) | Confirmed (rank 1) | | | Confirmed (rank 1) | | |

*Note: *** p<0.01; ** p<0.05; * p<0.10; n.s. = not significant. Newey-West correction applied. FMOLS handles endogeneity and serial correlation in cointegrated systems.*

**Table A4.2: Long-Run FMOLS Full Model Comparison**

| Variable | M1.a | M1.b | M2.a | M2.b | M3.a | M3.b | Notes |
|---|---|---|---|---|---|---|---|
| ln(M0) | 0.51*** | — | 0.35*** | — | 0.24*** | — | M0 specification |
| Inflation Rate | — | -1.57*** | — | -1.11** | — | 0.91** | Inflation specification |
| Reserves/Imports | 0.06 | -0.82*** | -0.11 | -0.71*** | -0.23** | -0.28* | |
| CA/GDP Ratio | -1.42 | 3.46*** | -1.39 | 4.06*** | — | — | Excl. in M3 |
| Policy Rate | -6.1* | 5.28* | — | — | — | — | Excl. in M2, M3 |

| Variable | M1.a | M1.b | M2.a | M2.b | M3.a | M3.b | Notes |
|---|---|---|---|---|---|---|---|
| Post-2024Q2 Dummy | -0.20 | -0.59*** | -0.49*** | -0.19*** | -0.38*** | -0.06 | |
| Constant | -2.38*** | 1.02*** | -1.77** | 1.24*** | -0.92** | 0.31* | |
| $R^2$ | 0.832 | 0.759 | 0.800 | 0.738 | 0.792 | 0.601 | |
| Adjusted $R^2$ | 0.810 | 0.726 | 0.779 | 0.710 | 0.776 | 0.570 | |

*Note: M1 includes policy rate (high multicollinearity, VIF >35). M2 is the preferred specification. M3 excludes CA/GDP. All models confirmed cointegrated via Johansen test.*

### Annex A5: Short-Run Estimation Results (OLS)

Short-run dynamics are estimated using OLS in first differences. The dependent variable is Δln(Premium). Sample: 2015Q3–2025Q4 (42 observations after adjustment for lagged terms). The reform dummy D_2024Q3_SINGLE equals 1 only in 2024Q3, capturing the immediate premium compression at unification.

**Table A5.1: Short-Run OLS Preferred Specifications (Models 2.a and 2.b)**

| Variable | Model 2.a Coeff. | Sig. | Model 2.b Coeff. | Sig. | Variable Definition |
|---|---|---|---|---|---|
| Δln(Premium)(t−1) | +0.23 | ** | +0.20 | ** | Lagged premium growth (momentum) |
| Δln(M0) | +0.55 | ** | — | — | Quarterly growth of monetary base |
| Δ Inflation Rate | — | — | +0.10 | n.s. | Change in domestic CPI inflation |
| Δ Reserves/Imports | +0.11 | n.s. | +0.05 | n.s. | Change in reserve adequacy |
| Δ CA/GDP Ratio | −0.55 | n.s. | −0.35 | n.s. | Change in current account/GDP |
| 2024Q3 Reform Dummy | −0.54 | *** | −0.50 | *** | Immediate FX unification shock |
| Constant | −0.01 | n.s. | +0.01 | n.s. | — |
| $R^2$ | 0.767 | | 0.733 | | |
| Adjusted $R^2$ | 0.735 | | 0.696 | | |

| Variable | Model 2.a Coeff. | Sig. | Model 2.b Coeff. | Sig. | Variable Definition |
|---|---|---|---|---|---|
| Heteroskedasticity (BP) | Homoskedastic | | Homoskedastic | | Breusch-Pagan-Godfrey |
| Serial Correlation (BG) | No serial corr. | | No serial corr. | | Breusch-Godfrey LM Test |
| Normality (Jarque-Bera) | Normal | | Not normal* | | *mild; robust to BP correction |
| CUSUM Stability | Stable | | Stable | | No structural breaks detected |

*Note: *** p<0.01; ** p<0.05; * p<0.10; n.s. = not significant. 42 observations. Standard errors robust to heteroskedasticity.*

**Table A5.2: Short-Run OLS Full Model Comparison**

| Variable | M1.a | M1.b | M2.a | M2.b | M3.a / M3.b | Notes |
|---|---|---|---|---|---|---|
| Δln(Premium)(t−1) | 0.15 | 0.11 | 0.23** | 0.20** | 0.20** / 0.18** | Momentum; sig. from M2 |
| Δln(M0) | 0.54** | — | 0.55** | — | 0.49** / — | Dominant SR driver |
| Δ Inflation Rate | — | 0.14 | — | 0.10 | / 0.04 | Insignificant in all SR models |
| Δ Reserves/Imports | 0.09 | 0.05 | 0.11 | 0.05 | 0.04 / 0.01 | Insignificant at QF |
| Δ CA/GDP Ratio | -0.48 | -0.29 | -0.55 | -0.35 | / — | Excl. in M3 |
| Δ Policy Rate | -1.93 | -2.30 | — | — | / — | Excl. M2+; high VIF in M1 |
| 2024Q3 Reform Dummy | -0.44*** | -0.38*** | -0.54*** | -0.50*** | -0.52*** / -0.49*** | Robust across all specs |
| Adjusted $R^2$ | 0.734 | 0.697 | 0.735 | 0.696 | 0.731 / 0.700 | |

*Note: M1 includes policy rate; M2 is preferred (no policy rate; parsimonious). Reform dummy highly significant and stable across all specifications.*

### Annex A6: Diagnostic Tests Summary

The table below summarises diagnostic test results for the preferred short-run models. All models pass standard regression diagnostics, confirming coefficient reliability.

| Test | Model 2.a (M0) | Model 2.b (Inflation) | Threshold | Interpretation |
|---|---|---|---|---|
| Heteroskedasticity (Breusch-Pagan) | Pass | Pass | $p > 0.05$ | Residuals homoskedastic; OLS efficient |
| Serial Correlation (Breusch-Godfrey) | Pass | Pass | $p > 0.05$ | No autocorrelation in residuals |
| Normality (Jarque-Bera) | Pass | Mild failure* | $p > 0.05$ | *Non-normality mild; robust to BP correction |
| Multicollinearity (VIF) | All < 3 | All < 3 | VIF < 10 | No problematic collinearity in M2 |
| CUSUM Stability Test | Stable | Stable | Within bounds | No structural breaks in coefficients |
| CUSUM-Square Test | Stable | Stable | Within bounds | No sudden parameter changes over time |

*Note: M1 models (with policy rate) had VIF > 7 on policy rate and reform dummy, motivating exclusion in M2. FMOLS long-run models do not require separate heteroskedasticity/serial correlation tests as the estimator corrects for both.*

### Annex A7: Variance Inflation Factors (VIF) Short-Run Models

VIF values confirm low multicollinearity in the preferred Models 2.a and 2.b, validating individual coefficient interpretation. Model 1 (with policy rate) exhibits high VIF on policy rate and reform dummy, motivating the exclusion of the policy rate in the preferred specification.

| Variable | M1.a VIF | M1.b VIF | M2.a VIF | M2.b VIF | Assessment |
|---|---|---|---|---|---|
| Δln(Premium)(t−1) | 2.23 | 2.18 | 1.13 | 1.11 | Low acceptable |
| Δln(M0) | 1.14 | — | 1.14 | — | Low acceptable |
| Δ Inflation Rate | — | 1.22 | — | 1.21 | Low acceptable |
| Δ Reserves/Imports | 2.60 | 2.58 | 2.55 | 2.56 | Moderate acceptable |
| Δ CA/GDP Ratio | 1.78 | 1.80 | 1.73 | 1.78 | Low acceptable |
| Δ Policy Rate | 7.01 | 7.06 | — | — | High excluded in M2 |

| Variable | M1.a VIF | M1.b VIF | M2.a VIF | M2.b VIF | Assessment |
|---|---|---|---|---|---|
| 2024Q3 Reform Dummy | 7.08 | 6.89 | 1.76 | 1.60 | High in M1, low in M2 |

*Note: VIF > 10 generally indicates problematic multicollinearity. Policy rate exclusion in M2 resolves the VIF issue and does not materially affect the other coefficients.*

### Annex A8: Model Specifications

**Long-Run Model (FMOLS) Preferred Specifications:**

Model 2.a: $\ln(\text{Premium}_t) = \alpha + 0.35 \cdot \ln(M0_t) - 0.11 \cdot (\text{Res/Imp})_t - 1.39 \cdot (\text{CA/GDP})_t - 0.49 \cdot \text{DPost} + u_t$

Model 2.b: $\ln(\text{Premium}_t) = \alpha - 1.11 \cdot \text{Inflation}_t - 0.71 \cdot (\text{Res/Imp})_t + 4.06 \cdot (\text{CA/GDP})_t - 0.19 \cdot \text{DPost} + u_t$

**Short-Run Model (OLS in First Differences) Preferred Specifications:**

Model 2.a: $\Delta\ln(\text{Premium}_t) = \alpha + 0.23 \cdot \Delta\ln(\text{Prem}_{t-1}) + 0.55 \cdot \Delta\ln(M0_t) + 0.11 \cdot \Delta(\text{Res/Imp})_t - 0.55 \cdot \Delta(\text{CA/GDP})_t - 0.54 \cdot \text{D2024Q3} + \varepsilon_t$

Model 2.b: $\Delta\ln(\text{Premium}_t) = \alpha + 0.20 \cdot \Delta\ln(\text{Prem}_{t-1}) + 0.10 \cdot \Delta\text{Inflation}_t + 0.05 \cdot \Delta(\text{Res/Imp})_t - 0.35 \cdot \Delta(\text{CA/GDP})_t - 0.50 \cdot \text{D2024Q3} + \varepsilon_t$

Where: DPost = dummy variable equal to 1 from 2024Q3 onward (post-unification regime); D2024Q3 = dummy variable equal to 1 only in 2024Q3 (immediate reform shock); all variables in natural logarithms or ratios as specified; Newey-West standard errors applied in LR; standard errors robust to heteroskedasticity in SR.

### Annex A9: Model Selection Rationale

| Criterion | Model 1 (Full) | Model 2 (Preferred) | Model 3 (Parsimonious) | Decision |
|---|---|---|---|---|
| Policy rate included | Yes | No | No | Excluded: VIF >7; theoretical overlap with M0 |
| CA/GDP included | Yes | Yes | No | Retained in M2: significant in LR Model 2.b |
| Multicollinearity (VIF) | High (>7) | Low (<3) | Low (<3) | M2 preferred on parsimony |
| Long-run fit (Adj. $R^2$) | 0.810 / 0.726 | 0.779 / 0.710 | 0.776 / 0.570 | M2 balances fit and parsimony |
| Short-run fit (Adj. $R^2$) | 0.734 / 0.697 | 0.735 / 0.696 | 0.731 / 0.700 | Comparable across models |

| Criterion | Model 1 (Full) | Model 2 (Preferred) | Model 3 (Parsimonious) | Decision |
|---|---|---|---|---|
| Normality | Pass / Pass | Pass / Mild fail | Pass / Pass | Mild non-normality in M2.b; BP-robust |
| Reform dummy significance | *** / *** | *** / *** | *** / *** | Robust across all specifications |

*Note: Model 2 is the preferred specification. It excludes the policy rate (high VIF with M0 and reform dummy) while retaining the current account variable that proves significant in the inflation model (2.b). Both sub-models (a = M0 specification; b = inflation specification) are reported for robustness.*

## Annex A10: Comparative Case Studies :Detailed Country Narratives

This annex provides the full case study narratives for Nigeria, Angola, and Argentina referenced in Section 4.1. Each case documents the reform context, post-reform parallel market premium dynamics, contributing factors, and lessons for Ethiopia.

**Case Selection:**

To situate Ethiopia's post-unification experience within a broader policy context, this section reviews recent exchange rate unification and liberalization episodes in Angola, Nigeria, and Argentina. All three countries undertook major FX regime changes with IMF engagement and experienced differing degrees of success in eliminating parallel market premia. Their experiences provide concrete evidence on how policy credibility, reserve buffers, capital controls, and monetary conditions shape post-reform outcomes.

IMF and central bank reports are used to track premium dynamics, policy sequencing, and institutional conditions across cases. The comparison focuses on three structural dimensions: remaining exchange restrictions, capital account openness, and financial market depth. Each case reflects country-specific circumstances, and the analysis does not imply a one-size-fits-all approach. The case studies are used diagnostically, to assess whether Ethiopia exhibits similar patterns and prescriptively, to identify policy measures that succeeded or failed in comparable settings.

## Nigeria: The Pitfalls of Half-Hearted Unification

**Context:** Nigeria has a long history of FX market fragmentation, with two key reform episodes in **2016–2017** and **2023–2024**, when authorities attempted to unify or float the naira.

- **2016 attempt:** Following an oil-price shock, Nigeria shifted from a fixed peg to a "managed float" in **June 2016**. However, the central bank (CBN) retained heavy management and multiple FX windows and did not fully liberalize the market.
- **2023 attempt:** the government formally scrapped the multiple-rate system and allowed the naira to depreciate by about **60 percent overnight**, aiming for unification under IMF monitoring

**Post-Reform Premium Dynamics:**

- After the **2016** partial float, the parallel premium narrowed from around **50 percent to ~20 percent** but never disappeared. By **2017**, the official rate was effectively re-pegged near, reopening a **~10 percent** gap.
- By late **2023**, the parallel rate weakened, reopening a **~15 percent** premium that persisted at roughly **~20 percent** into early **2024**. Continued intervention, FX supply restrictions amid low reserves, and expectations of further depreciation contributed to the relapse. High nominal interest rates (around **18 percent**) failed to offset inflation exceeding **20 percent**, sustaining dollar demand.

**Factors and Lessons:**

- **Credibility and Consistency:** In **2016**, mixed signals and multiple windows undermined credibility. In **2023**, clearer initial messaging was later diluted by policy reversals and ad hoc interventions.
- **Reserve Adequacy:** Reserves were weak in both episodes (around **US$33bn** in **2023**), limiting the CBN's ability to sustain a free float and forcing de facto rationing.
- **Inflation and Interest Rates:** Persistently high inflation (**15–20 percent**, rising above **20 percent** post-2023) meant real rates were weak or negative, encouraging conversion into FX and fueling parallel demand.

- **Political Economy:** Fuel-subsidy removal in **2023** (with prices quadrupling) triggered political backlash, leading to partial reversals that eroded reform credibility.

**Summary:** Nigeria's experience shows that half-measures do not work. Exchange rate unification must be backed by credible commitment, adequate reserves, sufficiently tight real interest rates, and consistent communication. Ethiopia's clearer reform framework under HGER is an advantage, but sustained consistency and political commitment are essential to avoid similar setbacks.

## Angola: Gradual Progress After a Premium Plateau

**Context:** Angola floated its currency (kwanza) in early 2018 as part of an IMF program to address a severe economic crisis (oil price crash). Pre-float, Angola had multiple rates and a parallel premium over 150 percent. The reform unified rates and depreciated the kwanza by ~40percent percent initially.

**Post-Reform Premium Dynamics:** After the float, Angola's premium *did not disappear completely*. It fell from triple digits to about 20-30 percent within a year. But then it **stubbornly hovered around 20 percent for another year**. It gradually eased to ~15 percent by late 2019, and eventually under 5 percent by mid-2021 (about 3 years post-reform). Why the slow convergence?

- **Remaining Restrictions:** Even after unification, Angola maintained certain exchange restrictions such as oil surrender requirement. Notably, limits on FX for certain invisibles and some **surrender requirements** for oil companies persisted initially. Once those were lifted, the premium showed marked decline.
- **Closed Capital Account:** Angola kept tight capital controls throughout. The IMF's comparative index shows Angola's capital account restrictions covered ~80 percent of transactions. This meant Angolans seeking to invest abroad continued to fuel a black market. Additionally, low confidence in local currency meant capital flight pressures lingered.
- **Monetary Policy & Interest Rates:** Angola's central bank was initially slow to tighten monetary policy. They were financing fiscal deficits which undermined the currency. By 2019, they raised interest rates significantly. That helped reduce inflation and stabilize

the kwanza, thereby narrowing the premium. Once they credibly squeezed liquidity, the parallel rate came closer.

- **Comparative Structural Factors:** Angola's financial market was shallow, and one state oil company (Sonangol) dominated FX supply. This concentration and lack of market infrastructure (sounds akin to Ethiopia's CBE dominance) meant price discovery took time.

**Lessons from Angola:** Ethiopia's situation in 2024–25 closely mirrors Angola 2018–19. Angola's premium was persistent because reforms were only partial at first and macro policy needed calibration. But once Angola completed critical reforms (removed last restrictions) and got inflation down, the parallel market virtually vanished. This suggests Ethiopia likely needs to implement *the complementary measures* (e.g the 2.5 percent commission, etc.) and maintain tight policies for a sustained period for the premium to fully go. Angola's case also shows the importance of **external factors**: as oil prices rose in 2021, Angola's exports grew and reserves recovered, which helped finish the job. Ethiopia might similarly benefit if, say, coffee/export prices increase or foreign aid picks up, but it can't bank on that as it should aim to create its own cushion.

## Argentina: Repeated FX Unification Episodes and the Limits of Incomplete Reform

**Context**: Argentina provides one of the most instructive comparative cases on FX unification, having undergone multiple episodes of capital control imposition, partial liberalization, and relapse. The peso cepo (currency clamp) was first imposed in 2011 under President Cristina Fernández de Kirchner, generating a parallel premium that peaked above 60 percent by 2015. President Macri lifted controls in December 2015, achieving rapid unification through a 40 percent devaluation. By 2018 a fiscal crisis forced reimposition of controls, and by 2023 the parallel premium exceeded 100 percent. President Milei undertook a new unification in April 2025, backed by a US$20 billion IMF programme.

First Unification (2015) - Initial Success:

- **Rapid Premium Compression**: A one-step 40 percent devaluation brought the official rate close to the parallel rate within weeks, eliminating the premium rapidly. Pass-through to

inflation was limited because the parallel rate had already been incorporated into price-setting.

- **Absence of Fiscal Anchor**: Despite early success, the reform lacked a credible fiscal consolidation. The government continued financing a large deficit through monetary expansion, depleting reserves and undermining credibility.
- **Policy Reversal**: By 2018 a balance-of-payments crisis forced a US$44 billion IMF programme, but insufficient reform led to another crisis. Controls were reimposed in 2019, demonstrating that exchange rate unification without fiscal discipline is not durable.
- **Credibility Collapse** (2019–2024): Under Presidents Fernández and Massa, fiscal deficits were monetized, excess liquidity accumulated, and the parallel premium widened to over 100 percent. The parallel market became the de facto price-setting reference, illustrating how official rate credibility collapses when monetary expansion is unconstrained.
- **Second Unification** (2025): President Milei's April 2025 unification was backed by zero monetary financing of the deficit, a fiscal primary surplus, and US$20 billion in IMF support. The parallel dollar fell 12 percent on announcement; the official and parallel rates converged rapidly. Country risk fell from 900 to 720 basis points within days.

**Outcome**: The 2025 unification succeeded in compressing the premium rapidly due to the presence of a credible fiscal and monetary anchor, the critical missing element in 2015. As of mid-2025, the premium remained near zero, though sustainability depends on maintaining fiscal discipline and reserve adequacy.

**Lessons for Ethiopia:**

- Fiscal Anchor is Non-Negotiable: Argentina's 2015 experience shows that exchange rate unification without fiscal consolidation collapses. Ethiopia's post-reform period must maintain monetary discipline and avoid deficit monetization to prevent a similar relapse.
- The Premium Reflects Credibility, Not Just Quantity: Argentina's premium was not caused by FX scarcity per se but by administrative bottlenecks and the loss of official rate credibility and consistent with the theoretical framing that access restrictions and credibility concerns, not mere supply shortfall, drive the parallel premium.

- Sequencing Matters: Argentina's 2025 success required fiscal surplus before lifting controls, not after. Ethiopia should similarly ensure monetary and fiscal conditions are durably anchored before removing remaining FX access restrictions.

Comparing all three cases, a pattern emerges: **the more comprehensive and credible the reform effort, the quicker the parallel premium fell** (Argentina showing both rapid initial success and subsequent relapse, Angola intermediate, Nigeria slow. The case studies thus reinforce the policy recommendations of this paper.