

Response to Referee 2

Thank you for your positive assessment of our work, and for the important issues raised in your report.

Issue 1: Is country X really comparable to country Y? We thank you for your compliment, and agree that the paper attempts to tackle this problem in an acceptable manner.

Issue 2: Is period J really comparable with period K? This is certainly an important point. We tried to address this issue in part by having controls in our regressions (e.g. Table 5), but as the referee points out the volatility of the shocks might be different in different regimes and different time periods. The comparison between the Bretton Woods and Gold Standard regimes is well-taken. To provide some additional evidence that our results are *not* driven simply by difference in the volatility of the shocks, we have performed a few additional exercises. (Key output is available at Rose's website.)

2.1. We estimate our main regression (see p.14 in the paper) for each regime separately:

$$CS_{it} = \alpha + \gamma'X_{it} + \varepsilon_{it} \quad \text{for each Regime}_{it} \text{ from } \{IT, ERT, MT, NQG\}$$

We then look at the RMSE for each regression to see whether the volatility of the shocks differs substantially. Using a specification that includes the same controls as in Table 5 (M2 growth, real GDP growth, terms of trade growth, openness, and the budget balance relative to GDP), we find the following residual RMSE.

Table B1: Residual RMSE for cumulative success regressions

Regime			
IT	MT	ERT	NQG
1.48	1.23	2.18	2.23

Although there is some cross-regime difference in the volatilities, the difference is not very large. IT has only about 30% lower RMSE than ERT. Furthermore MT regimes have the lowest RMSE. We thus agree with the view that, as the referee suggests, differences in the volatilities of shocks are unlikely to be the main drivers of our results.

2.2. Suppose that volatility does vary predominantly over time (as in the Bretton Woods vs. Gold standard example). In this case, a simple thing to do is just to restrict the sample to the post-1990s period. We know that ERT and MT regimes dominated in the 1970s and 1980s; as the referee points out, their poor record (in hitting low inflation) might be due to the nature of shocks during these periods rather than the potential of the regimes *per se*. By restricting the sample to the post-1990 period we have a time span that is the same for every regime. We do not report the results from this estimation, but the reason is that they are virtually unchanged from what we have reported in the paper. Hence using only the time period when all regimes were in existence does not alter our conclusions. This further bolsters our confidence in the basic result.

2.3. There is a third, more minor point, which pertains to the forecasting exercise in Table 7. One *could* make the following argument concerning volatility: Suppose that ERT is as successful as IT in hitting inflation at low numbers *on average*. However, if shocks in the ERT regimes were more volatile, we might see that the central bank gets outside of the “success” zone more often even though on average it is in the success zone. For example, a sequence of inflation rates in an ERT regime could hypothetically be 5%, -1%, 5%, -1% with a mean of 2%; the mean is within the success zone even though each observation is outside of the success zone. The comparable sequence for an IT regime might be 1%, 3%, 1%, 3%, where both each observation and the mean are within the success zone.

To check whether these kinds of realizations are behind the results of our forecasting exercise, we constructed the 5-year average inflation for each country and period. We then ask: “Does the regime have any forecasting power at a horizon of 3, 5, or 7 years for the success rate in having the 5-year average inflation rate within our success band?” If we had applied this transformation to the artificial example above, ERT and IT would look the same, although their volatilities are different. In reality, however, the IT regime continues to be the best predictor for having success in hitting low inflation. We take this result as additional (marginal) evidence to conclude that differences in volatility of shocks are unlikely to be the main driver behind the results in our paper.

Issue 3: Is keeping inflation between 0 and 4% (or 1-3%) the appropriate measure of regime performance?

There is not much we can do here. We have taken inflation as the main measure of success because there is unanimous consensus that in the long run this is what monetary policy controls. It would be nice to study also volatility performance (which we have done to a certain degree in Fatás, Mihov and Rose, 2007), but there are issues in applying our earlier methods to the study of duration. Since inflation targeting started in 1990, there has only been one (or perhaps two) business cycle observed in most inflation targeters. So we’d simply have too few plausible observations to compare the efficacy of different monetary regimes in smoothing business cycle fluctuations.