The Effect of Inflation Targeting: A Mean-Reverting Mirage?

Petra M. Geraats*

University of Cambridge

August 2012

WORK IN PROGRESS
COMMENTS WELCOME

Abstract

Inflation targeting has become a popular monetary policy strategy during the last two decades. This has given rise to a lively debate about the empirical effects of the adoption of inflation targeting. Some influential empirical studies have argued that the apparent improved performance of inflation targeters is merely regression to the mean, and controlling for the initial condition they conclude that inflation targeting does not matter. This paper challenges these findings that the apparent benefits of inflation targeting have basically been a mean-reverting mirage. It finds that these tests of the effect of inflation targeting have low power. It is shown analytically how they could fail to find any effect even if inflation targeting has in fact been highly effective. The low power of the tests is further illustrated using simulation results. As a result, prominent empirical findings that inflation targeting does not matter due to regression to the mean are misleading as the tests lack power to distinguish an oasis from a mirage.

KEY WORDS: monetary policy; inflation targeting. JEL CODES: E52, E31, C52

*Faculty of Economics, University of Cambridge, Cambridge, CB3 9DD, United Kingdom. Email: Petra.Geraats@econ.cam.ac.uk.

1 Introduction

Inflation targeting has become a popular monetary policy strategy during the last two decades. This has given rise to a lively debate about the empirical effects of the adoption of inflation targeting. In a highly influential paper, Ball and Sheridan (2003) argue that the apparent improved performance of inflation targeters is merely 'regression to the mean'. They use a difference-in-differences specification that includes the initial condition to control for this, and conclude that inflation targeting does not matter. Ball (2010) uses a similar specification in the Handbook of Monetary Economics and also finds little evidence that inflation targeting has been beneficial.

This paper challenges these findings that the apparent benefits of inflation targeting have basically been a mean-reverting mirage. It finds that tests of the effects of inflation targeting using the Ball-Sheridan specification have low power. It shows analytically how this specification could fail to find any effect of inflation targeting even if it has been highly effective. It explores how the interpretation of the coefficients depends on the data generating process. In particular, the coefficient estimate of the inflation-targeting indicator variable in the Ball-Sheridan specification may not capture the difference due to inflation targeting, but rather the difference in performance post-inflation targeting. So, if inflation targeting is actually effective at reducing inflation, but to an average level similar to others, then the Ball-Sheridan specification could give the incorrect impression that inflation targeting has been ineffective.

In addition to showing analytically that estimates of the effect of inflation targeting using the Ball-Sheridan specification tend to be biased, the low power of tests based on it is illustrated using Monte Carlo simulations. For instance, for plausible parameter values the paper finds that there would be no evidence of a significant effect of inflation targeting in 63% of replications even if there has in fact been a statistically and economically significant reduction in inflation of 2 percent point. Thus, tests of the effect of inflation targeting based on the Ball-Sheridan specification tend to be unreliable.

The remainder of this paper is organized as follows. Section 2 sets up the framework for the analysis and provides a simple illustrative example that shows how the Ball-Sheridan specification could yield misleading results. This example is generalized to allow for persistence in section 3. The issue of sample selection and regression to the mean is considered in section 4 and section **??** concludes.

2 Analysis

The effect of inflation targeting could be estimated using a differences-in-differences approach by comparing the change in a variable X_i (e.g. inflation in country i) before and after the adoption of inflation targeting (the ‘treatment’) to the change in X_i for others (the ‘control group’). This leads to the specification

$$\Delta X_i = a + bI_i + \varepsilon_i \tag{1}$$

where $\Delta X_i \equiv X_{i2} - X_{i1}$ denotes the change in X_i from period 1 to 2, I_i is an indicator variable for inflation targeting in country i in period 2, ε_i is i.i.d. white noise. The coefficient a captures the average change in X in the control group, and b the effect of the treatment of inflation targeting on X.

However, suppose that countries with higher initial inflation are more likely to adopt inflation targeting (as is observed empirically), so that X_{i1} and I_i are positively correlated. In particular, Ball and Sheridan (2003) argue that X_{i1} may be high because of temporary shocks. If countries with high X_{i1} decide to adopt inflation targeting, then X_{i2} would be expected to be lower because of ‘regression to the mean’, even if inflation targeting were completely ineffective. So, estimation of (1) using ordinary least squares (OLS) would lead to a downward bias in b because of a negative correlation between I_i and ε_i, and thereby overestimate the reduction in X due to the treatment effect.

To overcome this problem, Ball and Sheridan (2003) suggest to include the initial condition X_{i1}, so

$$\Delta X_i = a + bI_i + cX_{i1} + \varepsilon_i \tag{2}$$

If there is regression to the mean for X, the coefficient c for the initial condition X_{i1} would be expected to be negative, so a higher initial value X_{i1} reduces ΔX_i, leading to a relatively lower level of X_{i2}. The coefficient b is meant to capture the effect of the treatment of inflation targeting on X, corrected for regression to the mean.

To better understand the properties of the Ball-Sheridan (BS) specification (2), we first consider a simple illustrative example.

2.1 Illustrative Example

Assume that X_{it} is described by

$$X_{it} = \begin{cases} \mu_{Ot} + \varepsilon_{it} & \text{for} \quad I_i = 0 \\ \mu_{It} + \varepsilon_{it} & \text{for} \quad I_i = 1 \end{cases} \tag{3}$$

where μ_{It} and μ_{Ot} denote the average level of X in period t with inflation targeting and without inflation targeting, respectively, and ε_{it} is i.i.d. white noise with $\mathrm{E}[\varepsilon_{it}] = 0$ and $\mathrm{Var}[\varepsilon_{it}] = \sigma_{it}^2 \geq 0$ for all i and t, so X_{i1} and X_{i2} are independent. Suppose that inflation targeting is effective at achieving the inflation target X^* on average in period 2 so that $\mu_{I2} = X^*$, while other countries have an average of $\mu_{O2} = \mu_O$. So,

$$X_{i2} = \begin{cases} \mu_O + \varepsilon_{it} & \text{for} \quad I_i = 0 \\ X^* + \varepsilon_{it} & \text{for} \quad I_i = 1 \end{cases} \tag{4}$$

Note that the BS specification (2) can also be written as

$$X_{i2} = a + bI_i + (1 + c) X_{i1} + \varepsilon_i \tag{5}$$

This means that

$$X_{i2} = \begin{cases} a + (1 + c) X_{i1} + \varepsilon_i & \text{for} \quad I_i = 0 \\ a + b + (1 + c) X_{i1} + \varepsilon_i & \text{for} \quad I_i = 1 \end{cases}$$

Matching coefficients with (4) yields $c = -1$ and $\varepsilon_i = \varepsilon_{it}$, as the result should hold for any realization of X_{i1} and ε_{it}. Focusing on $I_i = 0$ and $I_i = 1$ then gives $a = \mu_O$ and $a + b = X^*$, respectively, which implies $b = X^* - \mu_O$. As a result, the BS specification (2) yields $a = \mu_O$, $b = X^* - \mu_O$ and $c = -1$.

This result also follows from the estimation of (5) by ordinary least squares (OLS). Let N be the number of observations in the sample, including $N_I \in \mathbb{N}$ inflation targeters and $N_O \in \mathbb{N}$ without inflation targeting in period 2, where $N = N_O + N_I$. The observations X_{it} are described by (3). For analytical convenience, assume that $\sum_{i \in R} \varepsilon_{it} = 0$ where R denotes the monetary policy regime (with $I_i = 0$ or $I_i = 1$), so the sample average $\bar{X}_{Rt}$ of X_{it} equals $\bar{X}_{Ot} = \mu_{Ot}$ and $\bar{X}_{It} = \mu_{It}$ for $I_i = 0$ and $I_i = 1$, respectively. Assume also that $\sum_{i \in R} \varepsilon_{i1} \varepsilon_{i2} = 0$, so the OLS estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta} \equiv (a, b, 1 + c)'$ satisfies $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}$ exactly.[1] Then the appendix shows that the OLS estimate for (5) equals

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \mu_{O2}, & \mu_{I2} - \mu_{O2}, & 0 \end{pmatrix}' \tag{6}$$

[1] This presumes that $N \geq 3$ and $\exists \varepsilon_{i1} \neq 0$ to ensure the three parameters in $\boldsymbol{\beta}$ can be estimated.

So, again $a = \mu_{O2} = \mu_O$, $b = \mu_{I2} - \mu_{O2} = X^* - \mu_O$ and $1 + c = 0$, so $c = -1$. The same outcome is obtained for OLS estimation of (2).

This result has important implications for the interpretation of the coefficients in the BS specification. When the data are described by (3), the intercept a equals the average period 2 level of X for countries in the control group without inflation targeting, rather than the average change in X in the control group. Furthermore, the coefficient b does not capture the *average change* in X due to the treatment of inflation targeting, but the *difference in the average level* of X between the treatment and control group in period 2. Finally, the variable X_{i1} capturing the initial condition has a negative coefficient with a magnitude of one, or a zero coefficient in the specification (5) in levels. The latter result is intuitive since X_{i1} and X_{i2} are assumed to be independent according to (3).

This illustrative example shows how the coefficients in the BS specification could be completely misinterpreted. In particular, consider the plausible case in which countries that adopted inflation targeting initially had a structurally higher level of inflation than others ($\mu_{I1} > \mu_{O1}$) and after the adoption of inflation targeting successfully reduced it to their inflation target which is set at $X^* = \mu_O$, whereas those without inflation targeting experienced no change in inflation ($\mu_{O1} = \mu_{O2} = \mu_O$). Then a regression using specification (2) would give a treatment coefficient $b = 0$, giving the incorrect impression that inflation targeting has been ineffective!

The same result holds if there was also a (smaller) decline in average inflation for those without inflation targeting, such that $\mu_{O1} > \mu_{O2} = \mu_O$. No matter how high inflation (μ_{I1}) initially was before inflation targeting, whenever the inflation target is set close to the average level of inflation of others ($X^* \approx \mu_O$), the estimated treatment effect is close to zero ($b \approx 0$), despite the fact that inflation targeting has successfully reduced inflation. Clearly, b cannot be interpreted as the average effect of inflation targeting in this case.

3 Persistence

The example above is based on the strong assumption that X_{it} is independent over time, which is not realistic when focusing on inflation or many other macroeconomic variables. In particular, although inflation targeters tend to show little inflation persistence, for other countries inflation tends to be quite persistent (Benati 2008). So it is important to allow for persistence in X, in particular X_{Ot}.

Before analyzing a more general case below, suppose now that X follows a random walk for countries without inflation targeting, so $X_{i2} = X_{i1} + \varepsilon_{i2}$ for $I_i = 0$. In particular, assume that X_{i1} is still given by (3) in period 1, but that now for period 2

$$X_{i2} = \begin{cases} \mu_{O1} + \varepsilon_{i1} + \varepsilon_{i2} & \text{for} \quad I_i = 0 \\ \mu_{I2} + \varepsilon_{i2} & \text{for} \quad I_i = 1 \end{cases} \tag{7}$$

where ε_{i2} is i.i.d. white noise. So, the effect of ε_{i1} is persistent for countries without inflation targeting, whereas inflation targeters manage to break with the past and are no longer affected by ε_{i1}. Assume again that $\sum_{i \in R} \varepsilon_{it} = 0$ and $\sum_{i \in R} \varepsilon_{i1}\varepsilon_{i2} = 0$, and denote $\sum_{i \in R} \varepsilon_{it}^2 = S_{Rt}$ and $S_t = S_{Ot} + S_{It}$, where R denotes the regime (with $I_i = 0$ or $I_i = 1$). Then the appendix shows that the OLS estimate for (5) equals

$$\hat{\boldsymbol{\beta}} = \left(\begin{array}{ccc} \frac{S_{I1}}{S_1}\mu_{O1}, & \mu_{I2} - \mu_{I1} + \frac{S_{I1}}{S_1}\left[\mu_{I1} - \mu_{O1}\right], & \frac{S_{O1}}{S_1} \end{array} \right)' \tag{8}$$

The interpretation of the estimated coefficients is again quite different from what may be expected for the BS specification. The intercept a does not capture the average change in X in the control group (which equals zero here), but a fraction S_{I1}/S_1 of μ_{O1}, where S_{I1} captures the volatility of the shocks in period 1 for countries that subsequently adopt inflation targeting, with $0 < S_{I1}/S_1 < 1$.[2] Furthermore, the coefficient b does not equal the average change in X due to the inflation targeting treatment, which is equal to $\mu_{I2} - \mu_{I1}$ in this case. Instead, if inflation targeting is effective at breaking with the past and reducing average inflation from $\mu_{I1} > \mu_{O1}$ to $\mu_{I2} < \mu_{I1}$, then the estimated 'treatment' coefficient is *smaller* in magnitude than the actual effect. Thus, the *estimated treatment effect is biased* again. Note that this bias is increasing in S_{I1}. So, if inflation targeters experienced relatively high initial volatility (which is plausible since they tend to be small open economies), the bias in the estimated treatment effect would be exacerbated. Finally, the estimate for the 'mean-reversion' coefficient c is equal to $S_{O1}/S_1 - 1 = -S_{I1}/S_1 < 0$, so its magnitude is also increasing in the initial volatility for inflation targeters.

The bias in the treatment effect makes it likely that OLS estimation of the BS specification would fail to find that inflation targeting has been effective. This can be illustrated by a Monte Carlo simulation. Suppose that X_{it} is inflation described by (7), where $\mu_{O1} = 2$ and $\mu_{I1} = 4 > \mu_{I2} = 2$, and ε_{it} is normally distributed, $\varepsilon_{it} \sim N\left(0, \sigma_{Rt}^2\right)$, with

[2] The strict inequalities presume that $\exists \varepsilon_{i1} \neq 0$ for each regime R.

$\sigma^2_{Ot} = \sigma^2_{It} = 1$ and $N_O = N_I = 10$, so $N = 20$.[3] Then the OLS estimates for (2) are $\hat{a} = 1.00$ (0.64), $\hat{b} = -1.00$ (0.73) and $\hat{c} = -0.50$ (0.27), based on 100,000 replications (with standard errors in parentheses). It is straightforward to check that these coefficient estimates are consistent with the analytical result in (8). The null hypothesis that inflation targeting is ineffective $H_0 : b = 0$ cannot be rejected in 73% of replications, despite the fact that inflation targeting successfully reduced average inflation by 2 compared to the control group. Clearly, the BS specification has low power to detect the effect of inflation targeting.

It is useful to check to what extent this is driven by the relatively large variance of the shocks, which gives a 95% confidence interval under inflation targeting of $X_{i1} \in [2, 6]$ and $X_{i2} \in [0, 4]$ for $I_i = 1$, making a reduction from 4 to 2 look insignificant. Now suppose instead that $\sigma^2_{Ot} = \sigma^2_{It} = 1/4$, so that the 95% confidence intervals under inflation targeting, $X_{i1} \in [3, 5]$ and $X_{i2} \in [1, 3]$ for $I_i = 1$, are much tighter and no longer overlap. Then the simulations yield $\hat{a} = 1.00$ (0.57), $\hat{b} = -1.00$ (0.59) and $\hat{c} = -0.50$ (0.27), so the coefficient estimates remain the same (in line with (8)) while the standard errors are reduced, but $H_0 : b = 0$ can still not be rejected in 63% of replications using the BS specification.

In sharp contrast, using the specification in differences (1) without the initial condition X_{i0} (i.e. restricting $c = 0$), OLS estimation yields the unbiased result $\hat{a} = 0.00$ (0.19) and $\hat{b} = -2.00$ (0.27), and rejects $H_0 : b = 0$ in all replications, using the same simulation. This is despite the fact that X_{i1} and I_i are highly correlated with a coefficient of 0.90. Clearly, a strong correlation between X_{i1} and I_i need not imply that OLS estimation of (1) is biased.

So far, the results in this section have been based on the assumption that X_{it} is independent over time for inflation targeters. However, it is probably optimistic to presume that inflation targeting allows for a complete break with the past, so it is important to also allow for some persistence for inflation targeters. Nevertheless, assuming a random walk for inflation is problematic under inflation targeting. First of all, from a theoretical perspective, an effective inflation targeter is able to achieve an inflation target X^* on average regardless of past shocks, so ε_{i1} should not have a permanent effect. Furthermore, empirical evidence (Benati 2008) shows that inflation persistence is very low for inflation targeters, which is inconsistent with a random walk. So, a more general specification is used to model persistence.

[3] Ball and Sheridan (2003) and Ball (2010) also use a sample size of 20 for their regressions.

Assume that X_{it} is still given by (3), except that now the assumption of independence between ε_{i1} and ε_{i2} is relaxed. Instead, let $\varepsilon_{i2} = \rho_R \varepsilon_{i1} + \eta_{i2}$, where ρ_R denotes the persistence parameter for regime R, with $0 \leq \rho_R \leq 1$, and η_{i2} is i.i.d. white noise. This means that

$$X_{i2} = \begin{cases} \mu_{O2} + \rho_O \varepsilon_{i1} + \eta_{i2} & \text{for} \quad I_i = 0 \\ \mu_{I2} + \rho_I \varepsilon_{i1} + \eta_{i2} & \text{for} \quad I_i = 1 \end{cases} \tag{9}$$

This convenient hybrid specification nests the previous two data generation processes. In particular, $\rho_O = \rho_I = 0$ gives (3), while $\mu_{O2} = \mu_{O1}$, $\rho_O = 1$ and $\rho_I = 0$ yields (7). Assume again that $\sum_{i \in R} \varepsilon_{i1} = 0$, $\sum_{i \in R} \varepsilon_{i1}^2 = S_{R1}$ and $S_1 = S_{O1} + S_{I1}$, as well as $\sum_{i \in R} \eta_{i2} = 0$ and $\sum_{i \in R} \varepsilon_{i1} \eta_{i2} = 0$, where R denotes the regime ($I_i = 0$ or $I_i = 1$). Then the appendix shows that the OLS estimate for (5) equals

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \mu_{O2} - \bar{\rho} \mu_{O1} \\ \mu_{I2} - \mu_{O2} - \bar{\rho} \left[\mu_{I1} - \mu_{O1} \right] \\ \bar{\rho} \end{pmatrix} \tag{10}$$

where $\bar{\rho} \equiv \frac{1}{S_1} \left(\rho_O S_{O1} + \rho_I S_{I1} \right)$ is a weighted average of ρ_R, with the weight S_{R1}/S_1 reflecting the relative initial volatility in regime R.

For the special case in which $\rho_O = \rho_I = 0$, $\bar{\rho} = 0$ so (10) reduces to (6), while for $\mu_{O2} = \mu_{O1}$, $\rho_O = 1$ and $\rho_I = 0$, it follows that $\bar{\rho} = S_{O1}/S_1$ and (10) is equal to (8). It is clear from (10) that the bias in the estimated intercept and treatment effect is not specific to these two cases but holds more generally. In particular, the intercept and treatment effect would be the average change in the control group $(\mu_{O2} - \mu_{O1})$ and the average change in the treatment group compared to the control group $(\mu_{I2} - \mu_{I1}) - (\mu_{O2} - \mu_{O1})$, respectively. So, in general there is a bias of $(1 - \bar{\rho}) \mu_{O1}$ for the intercept, which is positive for $\mu_{O1} > 0$, and a bias of $(1 - \bar{\rho}) (\mu_{I1} - \mu_{O1})$ for the treatment effect, which is also positive for $\mu_{I1} > \mu_{O1}$. So, when inflation targeters initially had higher inflation on average than others, as has been the case in practice, but then managed to reduce it, the magnitude of the estimated treatment effect is biased downward. As a result, the BS specification underestimates the reduction in inflation under inflation targeting, making it likely to incorrectly conclude that inflation targeting has been ineffective.

There is only one special case in which the estimates for the intercept and treatment effect are unbiased in the BS specification: $\bar{\rho} = 1$, which requires $\rho_O = \rho_I = 1$, so X follows a random walk for both inflation targeters and others. But, as mentioned before, a random walk in inflation is incompatible with a successful inflation targeter who manages to break with the past and achieve an inflation target X^* on average. Therefore, if inflation targeting is indeed effective, then $\bar{\rho} \neq 1$ and the estimated treatment effect of inflation targeting using the BS specification is biased, making it less likely to find a reduction in inflation.

Note that this bias in the estimated treatment effect is due to the BS specification that includes the initial condition X_{i1} as explanatory variable in an attempt to control for regression to the mean. In the specification in differences (1) without the initial condition (i.e. restricting $c = 0$), there is no bias and the OLS estimates for a and b are $\left(\bar{X}_{O2} - \bar{X}_{O1}\right) = (\mu_{O2} - \mu_{O1})$ and $\left(\bar{X}_{I2} - \bar{X}_{I1}\right) - \left(\bar{X}_{O2} - \bar{X}_{O1}\right) = (\mu_{I2} - \mu_{I1}) - (\mu_{O2} - \mu_{O1})$, respectively.[4]

4 Selection

The analysis so far has allowed for initial differences between inflation targeters and others, such as $\mu_{I1} > \mu_{O1}$ and $S_{I1} > S_{O1}$, but it has not considered sample selection effects. Suppose now that X is still described by (**??**), but that inflation targeting is not effective and that there is no fundamental difference between inflation targeters and others, so $\mu_{Ot} = \mu_{It} = \mu_t$ and $\sigma^2_{Ot} = \sigma^2_{It} = \sigma^2_t$. Instead, countries that happen to have high inflation in period 1 with $X_{i1} > \mu_1$ decide to adopt inflation targeting, whereas others do not. This selection into inflation targeting means that $\bar{X}_{I1} > \mu_1 > \bar{X}_{O1}$. In period 2, however, $\bar{X}_{I2} = \bar{X}_{O2} = \mu_2$. Then OLS regression of the plain specification in differences (1) yields an estimate for b of $\left(\bar{X}_{I2} - \bar{X}_{I1}\right) - \left(\bar{X}_{O2} - \bar{X}_{O1}\right) = \bar{X}_{O1} - \bar{X}_{I1} < 0$, suggesting that inflation targeting has reduced X, although the true treatment effect is zero since inflation targeting has been assumed to be ineffective with $\mu_{Ot} = \mu_{It}$. Clearly, the estimated treatment effect is biased; the reduction in X is simply regression to the mean. The bias is caused by the violation of the assumption that $\sum_{i \in R} \varepsilon_{i1} = 0$ as $\sum_{i \in I} \varepsilon_{i1} > 0 > \sum_{i \in O} \varepsilon_{i1}$ due to sample selection, so that $\bar{X}_{I1} > \mu_{I1} = \mu_{O1} > \bar{X}_{O1}$. The BS specification includes the initial condition X_{i1} in an attempt to control for such a selection effect.

Note that the presence of regression to the mean presumes that the effect of ε_{i1} is temporary, so it does not apply if X follows a random walk so that the shock ε_{i1} has a

[4] This is also derived in the appendix.

permanent effect. Similarly, there cannot be regression to the mean if the higher level of X_{i1} for $I_i = 1$ is a structural feature due to a higher mean μ_{I1}. In the latter case, the plain regression in differences (1) yields an unbiased estimate of the treatment effect b, whereas the magnitude of b is generally biased downwards in the BS specification (2), as shown in section 3.

The crucial question is whether X_{i1} and I_i are correlated because of temporary shocks ε_{i1} or fundamental factors μ_{R1}. This may be hard to distinguish and it is likely to depend on the context. For instance, if X_{it} is the rate of inflation in one year, then a high level of X_{it} could plausibly be due to a temporary positive shock ε_{it}. But, if X_{it} is the (average) rate of inflation over a period of half a decade, then it is more likely to reflect a high structural factor μ_t. In the latter case, one would not expect X_{it} to go down due to regression to the mean.[5] In particular, some countries may have structural features (e.g. small open economy, weak institutions) that make it more difficult to control inflation. They may suffer from structurally high inflation that is unlikely to subside unless measures are taken to mitigate the problem in some way (e.g. inflation targeting).

In Ball and Sheridan (2003), the pre-targeting sample period is at least 5 years and even up to 30 years. So, X_{i1} is a longer run average that is unlikely to exhibit regression to the mean. This means that $\bar{X}_{I1} > \bar{X}_{O1}$ is mostly due to $\mu_{I1} > \mu_{O1}$. If countries with high μ_1 decide to adopt inflation targeting, X_{i1} and I_i are correlated, but OLS estimation of (1) is unbiased, whereas the BS specification (2) is biased, unless $\rho_I = \rho_O = 1$ (as shown in section 3). But in the latter case, X_{it} follows a random walk and the effect of ε_{i1} is permanent, so there cannot be regression to the mean, which was the motivation for the BS specification.

Nevertheless, that leaves the question to what extent the BS specification (2) outperforms (1) when sample selection based on ε_{i1} is an issue. This is considered in the following Monte Carlo simulation. Suppose that X_{it} follows a random walk during the pre-targeting period, which lasts 5 years, starting from an identical initial condition $X_{i0} = \bar{X}_0 = 2$ for all i. Let $\bar{X}_1$ be the sample mean during the pre-targeting period across all observations. Assume that country i decides to adopt inflation targeting if $\bar{X}_{i1} > \bar{X}_1$, and that inflation targeting is successful in breaking with the past and re-

[5]Following Ball and Sheridan's (2003) baseball analogy, if I have a low batting average in a few games, one may think it is just temporary (perhaps due to an injury). But if my low batting average persists over time, the problem is more likely to be structural (e.g. reflecting my lack of skill), so one would not expect my batting average to go up.

ducing X_{it} to an average level of $\bar{X}_{I2} = \bar{X}_0$, where X_{it} is described by (3) during the post-targeting period, which also lasts 5 years. All other countries with $\bar{X}_{i1} \leq \bar{X}_1$ continue to follow a random walk during the post-targeting period. The volatility of shocks is given by $\sigma^2_{it} = 1$ for all i and t, and the sample size $N = 20$. Then OLS estimation of (2) gives $\hat{a} = -0.167$ (0.548), $\hat{b} = -0.001$ (1.052) and $\hat{c} = -0.316$ (0.374), based on 100,000 replications (with standard errors in parentheses). The null hypothesis that inflation targeting is ineffective, $H_0 : b = 0$, cannot be rejected in 93% of replications. So, it would be extremely unlikely that the BS specification would find inflation targeting to be effective, despite reducing average inflation from $\bar{X}_{I1} > \bar{X}_1$ to $\bar{X}_{I2} = \bar{X}_0 = 2$ by an average amount of -1.2, and making inflation more stable and less persistent, moving from $\rho = 1$ to $\rho_I = 0$.

Interestingly, the plain specification in differences (1), which is known to biased due to the sample selection based on ε_{i1}, actually performs noticeably better than (2), although it still cannot reject $H_0 : b = 0$ in 75% of replications. So, its power is quite low in this case, but the BS specification is even worse. Clearly, the presence of sample selection based on ε_{it} does not mean that the BS specification will be more suitable.

5 Conclusion

In influential contributions to the literature on the empirical effects of inflation targeting, Ball and Sheridan (2003) and Ball (2010) suggest that apparent improvements such as a reduction in inflation simply reflect regression to the mean after countries with temporarily high inflation decide to adopt inflation targeting. Using a modified difference-in-differences specification that aims to control for this by including the initial condition, they find little evidence that inflation targeting has been beneficial.

This paper exposes the shortcomings of their empirical approach by showing that tests of the effect of inflation targeting based on their specification have low power. It finds that the interpretation of the coefficients of their specification depends on the data generating process. In particular, if the persistence in X_{it} is sufficiently small, the coefficient estimate of the inflation-targeting indicator variable in their regression does not capture the 'treatment effect' of inflation targeting, but rather the difference in performance post-inflation targeting. So, when inflation targeting has succeeded in reducing inflation to the level of others, the Ball-Sheridan specification suggests it has been ineffective. Furthermore, it is shown analytically that the estimated 'treatment effect' tends to be biased in their specification.

Some Monte Carlo simulations are used to illustrate the low power of tests based on the Ball-Sheridan specification. For instance, if inflation targeters manage to significantly reduce average inflation by two percent points, while inflation of others follows a random walk, there is no evidence of a significant effect of inflation targeting in 63% of replications. In addition, even in the presence of sample selection, the Ball-Sheridan specification may perform considerably worse than the traditional difference-in-differences regression.

To conclude, influential empirical findings that inflation targeting does not matter due to regression to the mean are misleading as their tests lack power to distinguish an oasis from a mirage.

6 Appendix

This appendix shows that estimation of $\boldsymbol{\beta} = (a, b, 1+c)'$ in (5) using OLS yields $\hat{\boldsymbol{\beta}} = \begin{pmatrix} \mu_{O2}, & \mu_{I2} - \mu_{O2}, & 0 \end{pmatrix}'$ when the sample satisfies (3) with $\sum_{i \in R} \varepsilon_{it} = 0$ and $\sum_{i \in R} \varepsilon_{i1}\varepsilon_{i2} = 0$, where R denotes the regime (with $I_i = 0$ or $I_i = 1$).

Without loss of generality, order the observations i such that $I_i = 0$ for $i = 1, ..., N_O$ and $I_i = 1$ for $i = N_O + 1, ..., N$. Then the $N \times 3$ matrix of observations and the $N \times 1$ vector of the dependent variable are given by

$$\mathbf{X} = \begin{bmatrix} \mathbf{1}_{N_O} & \mathbf{0}_{N_O} & \mu_{O1}\mathbf{1}_{N_O} + \boldsymbol{\varepsilon}_{O1} \\ \mathbf{1}_{N_I} & \mathbf{1}_{N_I} & \mu_{I1}\mathbf{1}_{N_I} + \boldsymbol{\varepsilon}_{I1} \end{bmatrix} \text{ and } \mathbf{y} = \begin{pmatrix} \mu_{O2}\mathbf{1}_{N_O} + \boldsymbol{\varepsilon}_{O2} \\ \mu_{I2}\mathbf{1}_{N_I} + \boldsymbol{\varepsilon}_{I2} \end{pmatrix}$$

where $\mathbf{1}_N \equiv (1, ..., 1)'$ and $\mathbf{0}_N \equiv (0, ..., 0)'$ are $N \times 1$ vectors of ones and zeros, respectively; and $\boldsymbol{\varepsilon}_{Ot} \equiv (\varepsilon_{1t}, ..., \varepsilon_{N_O t})'$ and $\boldsymbol{\varepsilon}_{It} \equiv (\varepsilon_{N_O+1,t}, ..., \varepsilon_{Nt})'$ are $N_O \times 1$ and $N_I \times 1$ vectors of ε_{it} for $I_i = 0$ and $I_i = 1$, respectively. Note that $\mathbf{1}'_{N_O}\boldsymbol{\varepsilon}_{Ot} = \sum_{i=1}^{N_O} \varepsilon_{it} = 0$ and $\mathbf{1}'_{N_I}\boldsymbol{\varepsilon}_{It} = \sum_{i=N_O+1}^{N} \varepsilon_{it} = 0$. In addition, $\boldsymbol{\varepsilon}'_{R1}\boldsymbol{\varepsilon}_{R2} = 0$ for $R \in \{O, I\}$. For ease of notation, let $\boldsymbol{\varepsilon}'_{Ot}\boldsymbol{\varepsilon}_{Ot} = \sum_{i=1}^{N_O} \varepsilon_{it}^2 = S_{Ot}$ and $\boldsymbol{\varepsilon}'_{It}\boldsymbol{\varepsilon}_{It} = \sum_{i=N_O+1}^{N} \varepsilon_{it}^2 = S_{It}$, and denote $S_t = S_{Ot} + S_{It}$ for $t \in \{1, 2\}$.

The OLS estimate equals $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. To compute $\hat{\boldsymbol{\beta}}$, start with straightforward matrix multiplication to get

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} N_O + N_I & N_I & N_O\mu_{O1} + N_I\mu_{I1} \\ N_I & N_I & N_I\mu_{I1} \\ N_O\mu_{O1} + N_I\mu_{I1} & N_I\mu_{I1} & N_O\mu_{O1}^2 + N_I\mu_{I1}^2 + S_1 \end{bmatrix}$$

using the fact that $\mathbf{1}'_{N_R}\mathbf{1}_{N_R} = N_R$, $\mathbf{1}'_{N_R}\boldsymbol{\varepsilon}_{R1} = 0$ and $\boldsymbol{\varepsilon}'_{R1}\boldsymbol{\varepsilon}_{R1} = S_{R1}$ for $R \in \{O, I\}$. Note that $\det(\mathbf{X}'\mathbf{X}) = N_O N_I S_1 > 0$, so $\mathbf{X}'\mathbf{X}$ is nonsingular. Its inverse equals

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{N_O N_I S_1} *$$

$$\begin{bmatrix} N_I\left(N_O\mu_{O1}^2 + S_1\right) & -N_I\left(N_O\mu_{O1}^2 + S_1 - N_O\mu_{O1}\mu_{I1}\right) & -N_O N_I \mu_{O1} \\ -N_I\left(N_O\mu_{O1}^2 + S_1 - N_O\mu_{O1}\mu_{I1}\right) & N_O N_I\left(\mu_{O1} - \mu_{I1}\right)^2 + N S_1 & -N_O N_I\left(\mu_{I1} - \mu_{O1}\right) \\ -N_O N_I \mu_{O1} & -N_O N_I\left(\mu_{I1} - \mu_{O1}\right) & N_O N_I \end{bmatrix}$$

Postmultiplying this expression by $\mathbf{X}'$ and simplifying gives

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \frac{1}{N_O N_I S_1} *$$
$$\left[\begin{array}{cc} N_I S_1 \mathbf{1}'_{N_O} - N_O N_I \mu_{O1} \boldsymbol{\varepsilon}'_{O1} & -N_O N_I \mu_{O1} \boldsymbol{\varepsilon}'_{I1} \\ -N_I S_1 \mathbf{1}'_{N_O} - N_O N_I \left(\mu_{I1} - \mu_{O1}\right) \boldsymbol{\varepsilon}'_{O1} & N_O S_1 \mathbf{1}'_{N_I} - N_O N_I \left(\mu_{I1} - \mu_{O1}\right) \boldsymbol{\varepsilon}'_{I1} \\ N_O N_I \boldsymbol{\varepsilon}'_{O1} & N_O N_I \boldsymbol{\varepsilon}'_{I1} \end{array}\right]$$

Using the fact that $\mathbf{1}'_{N_R}\boldsymbol{\varepsilon}_{R2} = 0$ and $\boldsymbol{\varepsilon}'_{R1}\boldsymbol{\varepsilon}_{R2} = 0$, postmultiplying by $\mathbf{y}$ yields

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \frac{1}{N_O N_I S_1}\left(\begin{array}{c} N_I S_1 N_O \mu_{O2} \\ -N_I S_1 N_O \mu_{O2} + N_O S_1 N_I \mu_{I2} \\ 0 \end{array}\right)$$

Therefore,

$$\hat{\boldsymbol{\beta}} = \left(\begin{array}{c} \mu_{O2} \\ \mu_{I2} - \mu_{O2} \\ 0 \end{array}\right)$$

Now assume instead that X_{i2} is given by (7), so that

$$\mathbf{y} = \left(\begin{array}{l} \mu_{O1}\mathbf{1}_{N_O} + \boldsymbol{\varepsilon}_{O1} + \boldsymbol{\varepsilon}_{O2} \\ \mu_{I2}\mathbf{1}_{N_I} \end{array}\right)$$

Then premultiplying (**??**) gives

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \frac{1}{N_O N_I S_1}\left(\begin{array}{c} N_I S_1 N_O \mu_{O1} - N_O N_I \mu_{O1} S_{O1} \\ -N_I S_1 N_O \mu_{O1} - N_O N_I \left(\mu_{I1} - \mu_{O1}\right) S_{O1} + N_O S_1 N_I \mu_{I2} \\ N_O N_I S_{O1} \end{array}\right)$$

Hence,

$$\hat{\boldsymbol{\beta}} = \left(\begin{array}{c} \frac{S_{I1}}{S_1}\mu_{O1} \\ \mu_{I2} - \left[\frac{S_{O1}}{S_1}\mu_{I1} + \frac{S_{I1}}{S_1}\mu_{O1}\right] \\ \frac{S_{O1}}{S_1} \end{array}\right) = \left(\begin{array}{c} \frac{S_{I1}}{S_1}\mu_{O1} \\ \mu_{I2} - \mu_{I1} + \frac{S_{I1}}{S_1}\left[\mu_{I1} - \mu_{O1}\right] \\ \frac{S_{O1}}{S_1} \end{array}\right)$$

Now assume instead that X_{i2} is given by (9), so that

$$\mathbf{y} = \left(\begin{array}{l} \mu_{O2}\mathbf{1}_{N_O} + \rho_O \boldsymbol{\varepsilon}_{O1} + \boldsymbol{\eta}_{O2} \\ \mu_{I2}\mathbf{1}_{N_I} + \rho_I \boldsymbol{\varepsilon}_{I1} + \boldsymbol{\eta}_{I2} \end{array}\right)$$

where $\boldsymbol{\eta}_{Ot} \equiv (\eta_{1t}, ..., \eta_{N_O t})'$ and $\boldsymbol{\eta}_{It} \equiv (\eta_{N_O+1,t}, ..., \eta_{Nt})'$ are $N_O \times 1$ and $N_I \times 1$ vectors of η_{it} for $I_i = 0$ and $I_i = 1$, respectively. Note that $\mathbf{1}'_{N_R}\boldsymbol{\eta}_{R2} = 0$ and $\boldsymbol{\varepsilon}'_{R1}\boldsymbol{\eta}_{R2} = 0$ for $R \in \{O, I\}$. Then premultiplying (**??**) gives

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \frac{1}{N_O N_I S_1} *$$
$$\begin{pmatrix} N_I S_1 N_O \mu_{O2} - N_O N_I \mu_{O1} \rho_O S_{O1} - N_O N_I \mu_{O1} \rho_I S_{I1} \\ -N_I S_1 N_O \mu_{O2} - N_O N_I (\mu_{I1} - \mu_{O1}) \rho_O S_{O1} + N_O S_1 N_I \mu_{I2} - N_O N_I (\mu_{I1} - \mu_{O1}) \rho_I S_{I1} \\ N_O N_I \rho_O S_{O1} + N_O N_I \rho_I S_{I1} \end{pmatrix}$$

Hence,

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \mu_{O2} - \frac{1}{S_1}(\rho_O S_{O1} + \rho_I S_{I1})\mu_{O1} \\ -\mu_{O2} + \mu_{I2} - \frac{1}{S_1}(\rho_O S_{O1} + \rho_I S_{I1})(\mu_{I1} - \mu_{O1}) \\ \frac{1}{S_1}(\rho_O S_{O1} + \rho_I S_{I1}) \end{pmatrix} = \begin{pmatrix} \mu_{O2} - \bar{\rho}\mu_{O1} \\ \mu_{I2} - \mu_{O2} - \bar{\rho}(\mu_{I1} - \mu_{O1}) \\ \bar{\rho} \end{pmatrix}$$

where $\bar{\rho} \equiv \frac{1}{S_1}(\rho_O S_{O1} + \rho_I S_{I1})$.

Consider now the plain regression in differences (1) without the initial condition X_{i1} (i.e. restricting $c = 0$). So

$$\mathbf{X}_D = \begin{pmatrix} \mathbf{1}_{N_O} & \mathbf{0}_{N_O} \\ \mathbf{1}_{N_I} & \mathbf{1}_{N_I} \end{pmatrix} \quad \text{and } \mathbf{y}_D = \begin{pmatrix} (\mu_{O2} - \mu_{O1})\mathbf{1}_{N_O} + (\rho_O - 1)\boldsymbol{\varepsilon}_{O1} + \boldsymbol{\eta}_{O2} \\ (\mu_{I2} - \mu_{O1})\mathbf{1}_{N_I} + (\rho_I - 1)\boldsymbol{\varepsilon}_{I1} + \boldsymbol{\eta}_{I2} \end{pmatrix}$$

Then

$$\mathbf{X}'_D\mathbf{X}_D = \begin{pmatrix} N_O + N_I & N_I \\ N_I & N_I \end{pmatrix} \quad \text{and } (\mathbf{X}'_D\mathbf{X}_D)^{-1} = \begin{pmatrix} \frac{1}{N_O} & -\frac{1}{N_O} \\ -\frac{1}{N_O} & \frac{N_O+N_I}{N_I N_O} \end{pmatrix}$$

while using $\mathbf{1}'_{N_R}\boldsymbol{\varepsilon}_{R1} = 0$ and $\mathbf{1}'_{N_R}\boldsymbol{\eta}_{R2} = 0$ gives

$$\boldsymbol{\mu}'_D\mathbf{y}_D = \begin{pmatrix} N_O(\mu_{O2} - \mu_{O1}) + N_I(\mu_{I2} - \mu_{I1}) \\ N_I(\mu_{I2} - \mu_{I1}) \end{pmatrix}$$

So, the OLS estimate of $\boldsymbol{\beta}_D \equiv (a, b)'$ in (1) is equal to

$$\hat{\boldsymbol{\beta}}_D = \begin{pmatrix} \mu_{O2} - \mu_{O1} \\ -(\mu_{O2} - \mu_{O1}) + (\mu_{I2} - \mu_{I1}) \end{pmatrix}$$