Vysoká škola ekonomická v Praze

Národohospodářská fakulta

Hlavní specializace: Ekonomická analýza

OUTPUT-INFLATION TRADE-OFF AFTER A QUARTER OF A CENTURY OF INFLATION TARGETING

diplomová práce

Autor: Martin Kamarád

Vedoucí práce: Ing. Pavel Potužák, Ph.D.

Rok: 2017

Prohlašuji na svou čest, že jsem diplomovou práci vypracoval samostatně a s použitím uvedené literatury.

Martin Kamarád

V Praze, dne 2. ledna 2017

Poděkování

Na tomto místě bych rád poděkoval vedoucímu diplomové práce, Ing. Pavlu Potužákovi, Ph.D. za ochotu a jeho podnětné komentáře.

Abstrakt

Práce podrobně zkoumá vliv přijetí režimu inflačního cílování na úroveň inflace, její variabilitu, výstup ekonomiky a jeho variabilitu v zemích, které explicitně přistoupily k cílování inflace. V rámci výzkumu využívám metodu propensity score matching, která je schopna vyřešit nedostatky, typicky se vyskytující v nenáhodných experimentech, jako je například self-selection, a díky tomu zprostředkovat nevychýlené odhady efektu explicitního cílování inflace. Využité modely zahrnují nearest neighbor, radius matching, kernel matching a inverse probability weighting.

Výsledky naznačují, že rozvinuté i rozvíjející se země cílující inflaci dosahují nižších úrovní inflace při vyšších úrovních růstu produktu v porovnání se zeměmi, které inflaci necílují. Obě skupiny zemí cílujících inflaci zároveň dosahují nižší variability inflace a nižší variability produktu než země, které inflaci necílují. S výjimkou variability inflace je nicméně většina odhadů výsledků statisticky nevýznamná.

Odhady výsledků jsou do určité míry závislé na zvolené matching metodě. Statisticky nejvýznamnější odhady zprostředkovává Radius matching s úzkými rádiy ($r=0.005$, $r=0.001$).

Vyvážení průměrných hodnot pozorovaných proměnných mezi kontrolní skupinou a skupinou zemí cílující inflaci se zdá být dostatečné a kvantitativně lepší než v případě předchozích výzkumů.

Klíčová slova: Propensity score matching, inflační cílování, trade-off výstupu a inflace, Phillipsova křivka

JEL klasifikace: E5, E42, C21

Abstract

This thesis estimates the treatment effect of inflation targeting adoption on inflation, inflation variability, output, and output variability for 25 explicit inflation targeting countries. I implement the propensity score matching methodology that takes into account the problems of non-experimental nature, such as selection bias or selection on observable, and allows me to effectively mimic properties of randomized experiment and compute unbiased treatment effect estimates. I introduce a variety of propensity score matching methods that were recently developed in the treatment effect literature, including *Nearest Neighbor*, *Radius* matching, *Kernel* matching, and *Inverse Probability Weighting*.

The results indicate that both industrial and developing inflation targeting countries exhibit lower inflation levels and at the same time higher output growth than non-targeting countries. The estimates are however in most cases statistically insignificant. Moreover, it appears that both industrial and developing countries achieve combination of lower inflation variability and output variability compared to non-targeting countries. Nonetheless, majority of the estimates are again statistically insignificant.

The results are to a small extent sensitive to the choice of propensity score matching method. Radius matching with tight calipers (*r=0.005, r=0.001*) tends to provide the most reliable estimates.

Balancing properties of the models are reasonable and compared to the previous research the standardised biases are quantitatively better.

Key words: Propensity score matching, Inflation targeting, Output-Inflation trade-off, Phillips curve

JEL classification: E5, E42, C21

Introduction

Inflation targeting (hereinafter "IT") as a monetary regime emerged globally after the failure of monetarism in the 1980s and after the collapse of fixed exchange rate pegs in the early 1990s. The emergence of IT quickly attracted attention of many researchers and policymakers. Even though the amount of work on the effectiveness of IT has increased substantially over the last decade, there appears to be no common consensus on the impact of this monetary policy on a country's macroeconomic performance. Most of the initial studies focused on the effect of the IT adoption on inflation and inflation variability. Several studies find that the IT is successful in reducing both level of inflation and its variability (Wu, 2004; Pétursson, 2004). Other studies however document that the IT adoption has no significant effect on either (Ball & Sheridan, 2003; Bernanke, et al., 1999; Neumann & von Hagen, 2002). Furthermore, Johnson (2002) suggests that IT adoption did not succeed in reducing the variability of inflation expectations. Mixed are also results on the effect on output and output variability. Ball & Sheridan (2003) find no evidence that IT reduces output volatility or increases output growth. Mishkin & Schmidt-Hebbel (2001) on the other hand argue that IT regime not only reduces output volatility but also lessens the sacrifice ratio.

It is important to note that most of the initial literature on this topic suffers from insufficient number of observations as well as from a low number of treatment subjects. As a result, the outcomes of the early studies are somewhat inconsistent. Moreover, majority of the initial research assess the causal effect of IT adoption by performing traditional time series or an event study analysis that compares the outcome variables before and after IT adoption through linear regression and differences-in-differences methods. In randomized controlled experiment, randomization ensures that, on average, treated subjects do not systematically differ from control subjects in both measured and unmeasured characteristics. The treatment effect can be therefore estimated directly by comparing outcomes of both groups. However, non-randomized experiments are often subject to selection bias in which treated subjects systematically differ from control group. The linear regression and differences-in-differences method may therefore provide biased outcome estimates (Austin & Stuart, 2015). Furthermore, it is important to note that the decision to target inflation is not random and it is in fact endogenous (Mishkin & Schmidt-Hebbel, 2001; Gertler 2005). We therefore face the *selection on observable* problem which renders common linear regression unreliable (Lin & Ye, 2007; Dehejia & Wahba, 2002; Heckman, Ichimura & Todd, 1997).

To take into account these fundamental issues, several studies (Lin & Ye, 2007, 2009, 2012; Lin 2010; Luccote, 2012; Ardakani, et al. 2015) recently assessed the effect of IT adoption using the average treatment effect literature and propensity score matching methods (hereinafter "PSM"). PSM can be used to eliminate problems of non-experimental nature, such as selection bias or selection on observable, and to effectively mimic properties of randomized experiment (Cochran & Chambers, 1965). Under correct arrangements the PSM can provide consistent and unbiased estimates of the average treatment effect (Rosenbaum & Rubin, 1983). Vega & Winkelried (2005) represent probably the first study on the topic of IT effectiveness that implemented the PSM methodology. The study finds that IT adoption succeeds in delivering lower inflation and inflation variability compared to non-IT countries. Similar results for developing countries were provided by Lin & Ye (2009, 2012). However, Lin & Ye (2007) find that the effects are rather small and statistically insignificant for industrial countries.[1]

By eliminating problems of non-experimental nature, the PSM methods should noticeably improve quality and reliability of the treatment effect assessment (Lin & Ye, 2007). Ardakani, et al. (2015) however points out that most of the PSM literature on this topic may suffer from model misspecification. Majority of the studies do not conduct sufficient robustness testing and do not at all assess balancing properties of the models. Quality and reliability of the estimated outcomes may therefore be questioned.

The aim of this thesis is to assess the effect of IT adoption on the relation between output and inflation, i.e. *output-inflation trade-off*, as well as the effect on relation between inflation variability and output variability by estimating the average treatment effect on treated (hereinafter "ATT"). To control for the self-selection problem of non-random IT adoption and to obtain unbiased estimates of the IT adoption effect I implement a variety of PSM methods recently developed in the treatment effect literature, including *Nearest Neighbor* and *Radius* matching, as well as non-parametric *Kernel* matching and semi-parametric *Inverse Probability Weighting*.

This thesis aims to improve the existing literature on the effectiveness of IT in four important ways. First, the implemented econometric methodology takes into account the self-selection

[1] This illustrates why it is important to distinguish between industrial and developing countries when assessing the effect of IT adoption. I further address this topic in section 2.2

problem, which may arise because a central bank's decision to adopt IT is endogenous (Mishkin & Schmidt-Hebbel, 2001; Gertler 2005). Not controlling for this would otherwise result in biased estimates. I implemented a variety of PSM methods, including non-parametric Kernel matching and semi-parametric Inverse Probability Weighting, which were only scarcely implemented in the previous literature.

Second, PSM is a two-stage process that first requires estimation of the probability of IT adoption i.e. of the *propensity score*. The propensity score is then used to conduct the matching itself. The previous PSM literature has broadly ignored the role of financial market development on the probability of IT adoption. This is in high contrast to theoretical literature that regards the level of financial market development as a crucial precondition for IT adoption (Ardakani, et al., 2015). That is why I estimate the probability of IT adoption also by including proxy variable for the level of domestic financial sector development (financial depth, i.e. the credit provided to private sector by domestic financial institutions as a percentage of GDP).

Third, developing and industrial countries might significantly differ in their institutional and other arrangements, and the IT adoption is therefore unlikely to have the same effect on both groups (Aizenman, et al., 2011). To control for this, I assign each of the countries to either developing or industrial dataset and assess the effect of IT adoption for each of the group separately.

Fourth, to provide evidence on the quality and reliability of the performed matching I perform robustness testing as well as analysis of balancing properties of the model which is very often missing in the previous PSM literature.

The thesis is organized as follows. Section 1 presents the most important theoretical contributions to the output-inflation trade-off topic and a brief introduction of the Taylor rule. Section 2 provides brief theoretical discussion over the main ways in which the IT regime can be described, i.e. as a rule and as a framework. Furthermore, I discuss why it is important to clearly distinguish between developing and industrial countries when assessing the effect of IT adoption. Section 3 presents in detail the topic-related literature that uses conventional methods to assess the IT adoption effect as well as studies that implement PSM methods. Section 4 is dedicated to the description of the datasets and of the variables used for estimation of propensity score and ATT. Section 5 discusses in detail reasons for introduction of PSM rather than of common linear regression. Moreover, the section describes the implemented PSM methodology and assumptions that must apply. I conduct the estimation of propensity scores itself at the end of the section. Section 6 provides a general overview of all matching methods that are used in this research to estimate the ATT. The ATT estimation is conducted in section 7, and the model is afterwards tested for robustness in section 8.

1 Monetary Policy & Output-Inflation Trade-off

Phillips (1958) was one of the first authors who described the relation between unemployment rate and inflation rate[2]. He assessed this relation on data for the United Kingdom and noticed that there appeared to be an inverse relation between the two. Similar patterns were soon found also in other countries and were translated into an explicit link by Samuelson & Solow (1960). During the 1960s and early 1970s, many economists were convinced that there existed a long-run Phillips curve which offered a trade-off between the level of inflation and the unemployment rate. However, attempts to exploit such a trade-off were self-defeating and lead to a period of stagflation in the 1970s (Walsh, 1998). That experience has convinced most policymakers that no such long-run trade-off exists (Walsh, 1998). Notwithstanding that, the debate over what, if any, trade-off exists remained. In the few following sections I briefly describe, the Phillips curve, the contributions on the topic from New Classical and New Keynesian economics, and the Taylor rule.

1.1 Phillips Curve

Following the works of Phillips (1958), Samuelson & Solow (1960) formulated an explicit negative link between the inflation rate and the unemployment rate, which is known as the Phillips curve. According to their results that were estimated from 25 years of data, it appeared that there are various combinations of choices between different degrees of unemployment and price stability (Samuelson & Solow, 1960). Keynesian economists soon accepted and incorporate this theory into their macroeconomic models. Many were convinced that the relation holds also in the long-run. However, already in the late 1970s stagflation struck, and it was clear that the described relationship must be more complicated (Walsh, 1998).

One of the first authors who criticized the idea of the long-run Phillips curve, i.e. the long run trade-off between inflation-rate and unemployment-rate, were Phelps (1967) and Friedman (1968).

Friedman (1968) presented a hypothesis of the natural rate of unemployment, which is defined as the unemployment-rate to which the economy tends to in the long-run. The natural rate of unemployment is determined by real factors. Fluctuations around the natural rate are result of unexpected changes of the overall price level that are caused by fluctuations of the

[2] In his case it was the rate of change of money wage rates instead.

aggregate demand. Both Phelps (1967) and Friedman (1968) concur that the long-run Phillips curve and also long-run aggregate supply curve are vertical and hence there is no long-run trade-off between inflation-rate and the unemployment-rate. Moreover, higher inflation rates are gradually incorporated into people's adaptive expectations[3]. Trying to sustain unemployment-rate lower than its natural-rate in the long-run might result in accelerating inflation (Friedman, 1968; Phelps, 1967).

1.2 Lucas & Rational Expectations

Lucas' model (1973) presents without doubt one of the major contributions to the New Classical school of economics. The model itself is built upon a classical presumption of perfect competition where the market equilibrium is always established as a result of wage and price flexibility. The nominal output is determined on the aggregate demand side of the economy and its fluctuations are caused by fluctuations of the money supply. While the previously mentioned presumptions do not particularly deviate from other models with classical foundations, the model was ground-breaking for at least two reasons. Not only did Lucas for the very first time incorporate the theory of rational expectations into a macro-economic model but he also managed to construct his model on rigorous micro-economic foundations. The aggregate supply function (hereinafter "AS") is derived from the utility function of a representative individual, which can be presented as a function of consumption and labour supply. Quantity supplied in each market is viewed as the product of a normal component common to all markets and a cyclical component which varies from market to market[4] (Lucas, 1973, p. 327)

$$y_t(z) = y_{nt} + y_{ct}(z) \tag{1.1}$$

---

[3] Agents form their expectations about future inflation based on their previous experience. If their expectations proved to be wrong and the agents suffer from "money illusion", they tend to learn from their previous mistakes and adjust their expectations accordingly (Friedman, 1968).

[4] For all variables the small caption represents logarithmic expression of the variable, e.g.: $y_t(z) = \ln(Y_t(z))$.

Where y_{nt} is the normal component, and y_{ct} represents the cyclical component[5]. The AS can be further expressed as (Lucas, 1973, p. 327)

$$y_t(z) = y_{nt} + \gamma[p_t(z) - E\big(p_t \big| I_t(z)\big)], \qquad \gamma > 0 \tag{1.2}$$

Where γ is the structural parameter of the cyclical component that reflects the technology and agents' preferences. Each market is operated by a single agent who sells her production for a price $p_t(z)$. The agent decides about the production not only based on her own price but specifically based on her relative price. However, the agent is unable to observe the overall price level in the economy (due to imperfect information) and therefore cannot distinguish between changes to relative and general price level. In response to that the agent incorporates a set of all available information $I_t(z)$ to form rational expectations about the future development of the overall price level, i.e.: $E\big(p_t \big| I_t(z)\big)$.With just a few adjustments the equation 1.2 can be formulated as (Lucas, 1973, p. 328)

$$y_t(z) = y_{nt} + \theta\gamma[p_t(z) - \overline{p_t}], \tag{1.3}$$

and aggregated for all markets to:

$$y_t = y_{nt} + \theta\gamma(p_t - \overline{p_t}) \tag{1.4}$$

The slope of the AS curve is determined by the parameter θ, for which $\theta \equiv \frac{\tau^2}{\sigma^2+\tau^2}$ applies (Lucas, 1973, p. 328). When the relative price variations τ^2 are small[6] the Phillips curve as well as the AS curve should tend to be almost vertical, because the individual prices are almost certain to reflect the overall price level change. In this case the agents will incorporate the price changes into their expectations about the overall price level and will not dramatically increase their production. On the other hand, when the general prices are stable and hence the price level variance σ^2 is low[7] the Phillips curve as well as the AS curve

[5] y_{nt} reflects the capital accumulation and population growth, and follows the trend line $y_{nt} = \alpha + \beta t$

[6] And hence $(\sigma^2 \gg \tau^2)$

[7] I.e. $(\tau^2 \gg \sigma^2)$

should tend to be almost horizontal. In this case the agents will interpret the price changes mostly as changes of their relative price and will respond with increased production (Lucas, 1973). The equation therefore implies that the reason for fluctuations of the real output from its natural level is the difference between the expected and the actual overall price level in the economy. In other words, an unexpected rise of the prices may be mistaken for a rise of agent's relative prices and lead to increase of the agent's production and in aggregate to increase of the overall output of the economy.

To make the model complete, Lucas (1973) implements the standard form of Aggregate demand (hereinafter "AD") which can be presented as (Lucas, 1973, p. 328):

$$y_t = m_t - p_t \tag{1.5}$$

The short-run equilibrium is afterwards obtained as an intersection of both AD and AS (Lucas, 1973, p. 329)

$$p_t = E(m_t) - y^* + \frac{1}{1+\gamma^\theta}[m_t - E(m_t)] \tag{1.6}$$

$$y_t = y^* + \frac{\theta}{1+\gamma^\theta}[m_t - E(m_t)] \tag{1.7}$$

We can see that the equilibrium output can be altered only by unexpected change to money supply. The magnitude of the reaction depends on the parameter θ, i.e. on the slope of the AS curve. However, this applies only for the short-run, in the long-run the agents incorporate the changes into their expectations and subsequent changes in the money supply are reflected only in the overall price level.

1.3 New Keynesian Economics

Compared to New Classical economics, New Keynesian economics emphasizes that fluctuations in employment and output are largely a results of fluctuations in nominal aggregate demand. The reason that nominal shocks have real effects is that nominal wages and prices are not fully flexible and change infrequently. Furthermore, large nominal rigidities are possible even if the frictions preventing full nominal flexibility are small (Ball, et al., 1988).

Ball, et al. (1988) incorporate the Keynesian theory into a model of imperfectly competitive firms that change their prices at discrete intervals in time, because price adjustments are subject of costs. The model provides results that are consistent with the Keynesian explanation for the Phillips curve and inconsistent with the Classical explanation (Ball, et al., 1988). Authors conduct testing of the theories by assessing the relationship between average inflation and the size of the real effects of nominal shocks by measuring the effect of nominal shocks by the slope of the short-run Phillips curve. The behaviour of a representative firm can be expressed as (Ball, et al., 1988, p. 21):

$$p_i^*(t) - p(t) = v[y(t) - \bar{y}(t)] + \theta_i(t), \qquad v > 0 \tag{1.8}$$

Where $p_i^* - p$ represents the individual real price under which the firm maximizes its profit, y is the aggregate output, $\bar{y}$ is the natural output level and θ_i is the specific idiosyncratic shock that the firm faces. A firm's profit therefore depends on the aggregate output y, its relative price $p_i - p$ and on the specific shock θ_i. The overall price level is the average of all prices. If we assume that y increases, the firm will produce more under the given relative price. If the price adjustments were costless, the firm would set $p_i = p_i^*$ at every opportunity. However, the model assumes that the prices are changed only at intervals of length λ and that every price change has a fixed cost F. Individual price p_i is then obtained as (Ball, et al., 1988, p. 22)

$$p_i = \frac{1}{\lambda}\int_{s=0}^{\lambda} E_t p_i^*(t+s)ds \tag{1.9}$$

i.e. the firm sets such price that averages its expected profit maximizing prices for the period (Ball, et al., 1988).

To complete the model, authors introduce log of exogenous nominal AD $x \equiv y + p$ that follows random walk with a drift (Ball, et al., 1988, p. 22)

$$x(t) = gt + \sigma_t W(t) \tag{1.10}$$

where $W(t)$ is a Wiener process. Combining all assumptions along with the one that the natural rate of output grows smoothly at rate μ, i.e. $\bar{y}(t) = \mu t$, and therefore that the average inflation equals $g - \mu$, the behaviour of the economy can be solved for as (Ball, et al., 1988, p. 23)

$$p(t) = (g - \mu)t + \int_{s=0}^{\infty} w(s;\lambda)dZ(t-s) \tag{1.11}$$

Where $(g - \mu)$ is the average inflation, $w(s;\lambda)$ is the nominal shock at $(t - s)$, and $dZ(t - s) \equiv \sigma_x dW(t-s)$ is the innovation in aggregate demand at $(t - s)$. Authors denote that the immediate effect of a shock to the price level is equal to zero due to the fact that an infinitesimal proportion of the firms change their prices immediately. Nonetheless the shock grows over time.

Behaviour of the real output follows the behaviour of the price level and can be described as (Ball, et al., 1988, p. 24)

$$y(t) - \bar{y}(t) = \int_{s=0}^{\infty} [1 - w(s;\lambda)]dZ(t-s) \tag{1.12}$$

Where $[1 - w(s;\lambda)]$ determines the size of the real effects of nominal shocks. Moreover, equation for the variance of output can be expressed as (Ball, et al., 1988, p. 24)

$$E\{[y(t) - \bar{y}(t)]^2 = \sigma_x^2 \int_{s=0}^{\infty} [1 - w(s;\lambda)]^2 ds \tag{1.13}$$

The study denotes that the variance of output depends on the variance of the demand shocks σ_x^2, and the size of the effect of shocks $[1 - w(s;\lambda)]$. Moreover, the equilibrium frequency of price changes can be expressed as (Ball, et al., 1988, p. 24)

$$\frac{\partial L(\lambda_i, \lambda^E)}{\partial \lambda_i}|_{\lambda_i = \lambda^E} = 0 \tag{1.14}$$

Authors find that λ is decreasing in $\bar{\pi}, \sigma_x$, and σ_θ, where $\bar{\pi}$ is the average inflation, i.e. $\bar{\pi} \equiv g - \mu$. Therefore, higher average inflation shortens the time interval between price adjustments. High inflation causes that nominal profit maximizing price changes more frequently and therefore increases the benefits of more frequent price adjustments. The interval between price adjustments also decreases the greater are the shock variations $(\sigma_x, \sigma_\theta)$, which is due to the fact that it is harder for a firm to forecast its future profit maximizing price. The firm therefore will not want to fix it for a long period of time (Ball, et al., 1988).

To summarize, Ball, et al. (1988) imply that the Phillips curve is steeper when $\bar{\pi}, \sigma_x$ or σ_θ is larger. Higher average inflation reduces the interval between price changes, increasing the proportion of a shock that is translated into prices, i.e. $w(s;\lambda)$. The same applies for aggregate or firm-specific shocks. Increase in $\bar{\pi}, \sigma_x$ or σ_θ therefore causes firms to adjust prices more frequently. In turn, more frequent price changes imply that prices adjust more quickly to nominal shocks and the shocks have smaller real effects, i.e. $[1 - w(s;\lambda)]$. Average inflation therefore affects the slope of the Phillips curve (Ball, et al., 1988).

1.4 Taylor Rule

Taylor (1979) defined and estimated a trade-off between the variability of inflation and variability of output that is consistent with rational expectations. As the author denotes, efforts to keep the inflation rate too stable would result in larger fluctuations in real GDP and unemployment. Conversely, efforts to smooth out the business cycle too much would result in a more variable inflation (Taylor, 1979). Following the discussion in section 1.1, most economists nowadays accept that there is no long-run trade-off between the inflation rate and the level of unemployment and at the same time that policies aimed at inflation stabilization do have real effects (Walsh, 1998). Taylor (1996) sets out two propositions and summarize that there is no long-run trade-off between the level of inflation and the rate of unemployment. However, there is a short-run trade-off between the variability of inflation and the variability of unemployment. Consider an adverse aggregate shock that causes a rise in inflation through inflation expectations. If a central bank acts to bring the inflation back to its target promptly, inflation will be less variable, but output will fluctuate more around the trend. On the contrary if central bank tolerates the inflation variation and brings it down slowly, then output will fluctuate less while inflation becomes more variable. Focus on inflation stabilization following an AS shock should therefore lead to increased output variation, whereas focus on output stabilization should increase inflation variation (Taylor, 1996; Walsh, 1998).

Back in the 1993 Taylor (1993) has captured and approximated responsiveness of nominal interest rate as set by the central bank to changes in inflation and output. Moreover, Taylor suggested that behaviour of the Federal Reserve System in the United States during 1987-1992 can be characterized by a rule that describes how the federal funds rate is adjusted in response to movements in inflation and the output gap (Taylor, 1993). Nowadays the rule is generally known as the *Taylor rule*. In particular, the Taylor rule describes how for each *X percentage* increase in inflation, the central bank tends to raise the nominal interest rate by more than *X percentage*. The rule is intended to enhance price stability by systematically increasing the credibility of future actions by the central bank (Taylor, 1996; Walsh, 1998). The Taylor's original specification that fitted the actual FED's policy performance during the 1987-1992 can be expressed as (Taylor, 1993, p. 202)

$$r = p + .5y + .5(p - 2) + 2 \tag{1.15}$$

Where r is the federal funds rate, p is the rate of inflation over the previous year, and y is the deviation of real GDP from a target[8]. The policy rule introduces such feature that the federal funds rate rises above $(2 + p)\%$ if the inflation rate increases above its target of 2% or if the real GDP rises above the trend GDP (Taylor, 1993).The policy can be further rewritten to the following general expression (Taylor, 1996, p. 192)

$$i = \pi + gy + h(\pi - \pi^*) + r^f \tag{1.16}$$

Where i is the target short-term nominal interest rate, measured in percentage points, π is the inflation rate, π^* is the target inflation rate, y measures the percentage deviation of real GDP from potential GDP, and r^f is the central bank's estimate of the equilibrium real interest rate. Coefficients g and h represents central bank's preference weights towards output and inflation deviation from their target. The coefficients are in all cases positive[9]. When inflation rises, the policy calls for interest rate increase by more than the inflation rate (Taylor, 1993). By changing how much the interest rate is adjusted in response to inflation and the output gap, a different combinations of inflation variability and output variability arise. If the different combinations of inflation and output variability are linked together we can observe a particular trade-off between the two (Taylor, 1993; Walsh, 1998).

---

[8] i.e. $y = 100\frac{(Y-Y^*)}{Y^*}$. Where Y is the real GDP, and Y^* is the trend real GDP.

[9] Taylor (1993) proposed a setting of .5 for both coefficients.

2 Inflation Targeting

IT as a monetary regime emerged globally after the failure of monetarism in the 1980s and after the collapse of fixed exchange rate pegs in the early 1990s. The major motivation for IT adoption was the fact that in most countries the relationship between intermediate targets, such as money growth or exchange rate, and the central bank’s goal variable has proven to be somewhat unreliable. Compared to aforementioned monetary regimes the IT framework significantly reduces the role of intermediary targets and instead attempts to target the goal variable directly (Bernanke & Mishkin, 1997). However, since the central bank is generally required to forecast the expected inflation development, it turns its focus typically to a variety of indicators that have proven their predictive power in the past (Bernanke & Mishkin, 1997).

It is widely accepted in the academic literature that the primary role for monetary policy is price stability (Bernanke, et al., 1999). Other objectives such as promoting growth and employment should be seen only as secondary objectives. IT is often being characterized as monetary policy of “constrained discretion” because it combines elements of both rules and discretion (Hammond, 2012). It provides a rule-like framework which can be used by the private sector to anchor its expectations about future inflation. However, central bank has discretion in reacting to shocks and in reaching its inflation target and achieving price stability (Hammond, 2012).

Before we continue, it is worth noting that while Mishkin & Schmidt-Hebbel (2001) agree that IT is a good framework for keeping inflation low and stable, and therefore for delivering price stability, it may not necessarily be the best monetary regime for bringing down high or even hyper-inflation. The ability to target inflation depends on quality of the forecasts, which are much less reliable when the inflation is high and volatile. Central bank is therefore likely to loose its credibility by getting the forecast wrong and having large target misses. Nontheless Mishkin & Schmidt-Hebbel (2001) note that for example Chile and Israel (and other) implemented the IT succesfully in the past even though their economies encountered abnormaly high inflation levels.

In the following sub-sections I provide brief theoretical discussion over the main ways in which the IT regime can be described, i.e. as a rule and as a framework. Furthermore, I discuss why it is important to clearly distinguish between developing and industrial countries when assessing the effect of IT adoption.

2.1 Brief Theoretical Background

Even after more than two decades some confusion persists on how to define IT. Kuttner (2004) suggests that there are two main ways to think about IT: first, in terms of the observed characteristics of the policy framework and the second in the terms of policy rule. The second way is also emphasized by Svensson (1996, 1999)[10]. However, Bernanke, et al. (1999) suggests that IT is better described as a framework that involves number of elements, rather than a rule.

2.1.1 IT as a Rule

Svensson (1996, 1999) argues that IT can be interpreted as an optimization rule that specifies a target variable and inflation target level to minimize an objective loss function over both inflation gap and output gap. Central bank's objective is to choose such interest rates (current and future) that minimize its loss function (Svensson, 1996, p. 612):

$$E_t \sum_{\tau=t}^{\infty} \delta^{\tau-t} L(\pi_\tau) \tag{2.1}$$

where π is inflation, E_t are the expectations conditional on information set available in year t, δ is the discount factor, and $L(\pi_\tau)$ is the loss function. According to Svensson (1999, p. 621) the loss function conventionally takes the form of:

$$L_t = \frac{1}{2}[(\pi_t - \hat{\pi})^2 + \lambda y_t^2] \tag{2.2}$$

where $\hat{\pi}$ is the inflation target level, y_t is the output gap and λ[11] is the relative weight that is applied on the output gap. The case when $\lambda = 0$ and only inflation gap enters the loss function is called "*strict inflation targeting*", whereas the case when $\lambda > 0$ and the output gap also enters the loss function is called "*flexible inflation targeting*" (Svensson, 1999). Under flexible IT the inflation is brought to its target more gradually compared to strict IT because central bank has to divide its focus on real-output variance as well on the inflation

---

[10] Woodford (2004) also describes IT in terms of policy optimization.

[11] $\lambda \geq 0$ has to apply

variance. Similarly, flexible IT should exhibit longer aim horizon for meeting the inflation target (Svensson, 1999).

Simply put, the rule expresses the need to balance the expected marginal benefit of reducing inflation (in form of deviation of π_t from $\hat{\pi}$) with the expected marginal cost of the inflation reduction (in form of negative y_t). A larger λ means that the inflation reduction comes at a greater cost, and as a result the optimising central bank will be willing to tolerate larger deviations of π_t from $\hat{\pi}$, hence the trade-off (Kuttner, 2004).

2.1.2 IT as a Framework

Based on practical observations, Truman (2003) argues that IT is more than just a rule and that IT framework involves a number of elements such as emphasis on transparency and communication towards public. These elements cannot be properly accounted for while deriving the policy rules (Truman, 2003).

Similarly, Bernanke & Mishkin (1997), Bernanke, et al. (1999), and Hammond (2012) provide practical examples that IT is not a policy rule, as is sometimes presented in the theoretical literature, but rather a policy framework, whose major advantage is high level of transparency and coherence policy in which the central bank can conduct flexible monetary policy actions. Miao (2009) argues that a closer look among IT countries reveals that this flexibility in monetary policy creates noticeable divergence among IT regimes that appears to be rather norm than an exception. Hammond (2012) compares IT frameworks of 27 IT countries and agrees that each of the individual frameworks reflects local economic, political and cultural factors that distinguishes it from the rest. However, there appears to be a wide common ground on the main features among each of the frameworks. Bernanke & Mishkin (1997) specify that the cornerstone of IT framework lies in the public announcement that the central bank will strive to keep inflation at some numerically specified level[12] and in the explicit acknowledgment that low and stable inflation, *i.e. price stability*[13] is the main goal for the monetary policy. Other important characteristics include increased transparency and communication towards public and also increased accountability of the central bank for attaining its objectives (Bernanke & Mishkin, 1997).

[12] i.e. establishment of official target or range for the inflation rate at one or more horizons

[13] Bernanke & Mishkin (1997) note that price stability is usually considered inflation close to a 2% annual rate of price change. Taylor (1996) mentions 1% or 2% annual rate of price change.

In practice we can see that the IT adoption is typically accompanied with a central bank's disclosure that control of inflation is the primary goal of monetary policy and that the central bank will be held accountable for meeting the inflation target. The IT adoption is also typically linked with changes to legislation or administrative arrangements in the direction of increased independence of the central bank such that the bank can be free to make the technical decisions necessary to achieve its goals. Furthermore, to reflect the objective of improved transparency and communication of the monetary policy and its goals towards public, most IT central banks regularly publishes detailed assessments of the inflation situation, including inflation forecasts (Bernanke & Mishkin, 1997).

2.2 Inflation Targeting in Emerging Countries

In this section I discuss why it is important to clearly distinguish between developing and industrial countries when assessing the effect of IT adoption. It is well documented that developing and industrial countries might significantly differ in their institutional arrangements, level of central bank's independency and other aspects. IT adoption is therefore unlikely to have the same effect on both groups (Aizenman, et al., 2011). Calvo & Mishkin (2003) specify six fundamental differences between institutional arrangements of industrial and developing countries that should be taken into account. The developing countries might suffer mainly from weak financial institutions, weak government prudential regulation and supervision, low credibility of monetary institutions, currency substitution, tendency for liability dollarization, and vulnerability to sudden economic stops. While industrial countries are not immune to these problems, the magnitude for developing countries is much more significant.

The aforementioned weaknesses make developing countries vulnerable to high inflation and currency crises. The real value of money often cannot be taken for granted and as a result domestic residents might be likely to conduct currency substitution. This might to what is called *liability dollarization* (Mishkin, 2004).

Masson, et al. (1997) and Mishkin (2004) specify several fiscal and monetary institutional prerequisites that the developing countries should meet in order to successfully operate IT regime.

2.2.1 Fiscal and Financial Institutions

According to Mishkin (2004) fiscal stability is one of the fundamentally necessary conditions for inflation control and hence successful operation of the IT regime. Masson, et al. (1997) stress that domestic monetary policy cannot be dictated or severely constrained by a country's problem with fiscal balance – i.e. *fiscal dominance*. Irresponsible fiscal policy and fiscal dominance puts pressure on the central bank to monetize the government debt, which might lead to a rapid money growth and high inflation. When the fiscal imbalances are high enough, the central bank might be rendered defenceless in attaining its inflation target.

Another crucial characteristic that might not be met in a developing country is a strong and deep banking and financial system. A weak financial system might be particularly dangerous in situations when the central bank is required to rapidly raise interest rates in order to sustain the inflation target. This might very well provoke collapse of the weak system, lead to a currency collapse, financial crisis and in the end cause breakdown of the IT regime itself (Mishkin, 2004). Moreover, Hammond, et al. (2009) argue that weak financial and banking system may also hamper the transmission mechanism. To avoid this potential collapse, developing countries should implement fiscal reforms that increase transparency of the governmental budget and implement rules that help keep the budget deficits under control[14] (Mishkin, 2004). Safe and sound financial system is therefore a necessary condition for the success of an IT regime (Mishkin, 2004; Masson, et al., 1997; Hammond, et al., 2009).

Mishkin (2004) further encourages implementation of policies aimed to increase trade openness of an economy. Since businesses in tradeable sector have often balance sheets denominated at least partly in foreign currencies (goods are likely to be priced in foreign currency) the businesses are less exposed to negative consequences from devaluation of the domestic currency. Devaluation may on one hand increase the value of their debt in terms of domestic currency, however it is likely to raise the value of their assets denominated in foreign currency (Mishkin, 2004).

It appears that fiscal and financial stability are necessary conditions for successful inflation control. However, it is not clear whether they should be viewed as prerequisites for IT adoption or whether they can be implemented gradually. Masson, et al. (1997) emphasize that such reforms should be in place before the IT adoption. However, Mishkin (2004) and

[14] According to Mishkin (2004) additional reforms should also focus on prudential regulation of the banking system and strengthening the financial system. Furthermore, a limit on safety assurance from the government should be introduced in order to control the moral hazard incentives for banks (Mishkin, 2004).

Bernanke, et al. (1999) note that because IT regime commits the monetary institution to keep inflation low, it can also help promote fiscal and financial reforms because it is clear that the government must support these reforms if the IT regime is to be successful (Mishkin, 2004; Bernanke, et al., 1999).

2.2.2 Monetary Institutions

According to Hammond, et al. (2009) central banks in developing countries face a unique set of challenges that might severely complicate the IT adoption. The major institutional deficiency is the lack of central bank independence. In some countries, this might take the form of statutory subordination under finance ministry. In other countries the central bank still might be controlled by political establishment. Hence, central banks in developing countries might have severe problems in terms of maintaining their legitimacy and independence (Hammond, et al., 2009). Moreover, central banks in developing countries face a number of technical challenges in implementing IT regime. The central bank needs the technical capacity to model the economy, understand the transmission mechanism and forecast inflation and output (Hammond, et al., 2009).

Mishkin (2004) further specifies that public and institutional commitment to price stability as well as commitment to independence of central bank are required in order for the monetary authority to operate IT regime successfully. The first one, i.e. institutional commitment to price stability, serves as the long-run goal of monetary policy, which grants central bank the mandate to control inflation. However, many developing countries have had a history of insufficient support for the price stability goal even after institutional commitment in form of laws[15]. Validity of such arrangements in developing countries might therefore be questionable (Mishkin, 2004). The second one, i.e. public and institutional commitment to independence of the central bank, means that the central bank has to be free to set the monetary policy instruments without any political pressure and also that central bank has to be prohibited from funding government (Mishkin, 2004). If a central bank is sufficiently independent, IT regime grants it some level of discretion and flexibility to cope with economic shocks and to attain its target goal. Fraga, et al. (2003) point out that the required level of policy flexibility and hence central bank independency is even higher for central banks in developing countries because they tend to be subject to larger and more

[15] See Mishkin (2004) for discussion over Latin and South America countries.

intensive economic shocks than industrial countries. Moreover, Masson, et al. (1997) argues that both requirements for commitment to price stability and for central bank independency might be hampered by shallow capital markets and fragile banking systems if the fiscal prerequisites discussed in the previous section are not met (Masson, et al., 1997).

3 Contemporaneous empirical research

Ardakani, et al. (2015) summarise three problems that the previous empirical literature on the effectiveness of IT suffers from: I. improper methodology for estimation of treatment effect II. the variables used to assess effect of IT adoption[16] ignore extant theoretical literature III. most of the research focuses primarily on inflation and inflation variability and omits other outcome variables.

Most of the initial research asses effect of IT adoption by performing time series or an event study analysis that compares the outcome variables before and after IT adoption through linear regression models[17]. However, if decision to target inflation is not random and it is in fact systematically correlated with a set of observables that also affect the outcomes, then we face the *selection on observable* problem which renders common linear regression unreliable (Lin & Ye, 2007; Dehejia & Wahba, 2002; Heckman, Ichimura & Todd, 1997). Recently, several studies started examining IT effectiveness using the average treatment effect literature and propensity score matching methods[18]. PSM can be used to eliminate problems of non-experimental nature, such as selection bias or selection on observable, and to effectively mimic properties of natural experiment (Cochran & Chambers, 1965). In this section I describe literature that uses conventional methods to assess IT adoption effect as well as studies that implement PSM methods.

3.1 Early Literature

Obvious shortcoming of the early literature is that it suffers from insufficient number of observations as well as low number of treatment subjects. We can often see that outcomes of the early studies conducted in the 1990s are somewhat inconsistent not only with more

[16] Including variables used to assess likelihood of IT adoption, i.e. propensity score estimation.

[17] For example, Groeneveld (1998), Bernanke, et al. (1999), Kuttner & Posen (1999), Mishkin (1999), Johnson (2002), Ball & Sheridan (2003).

[18] For example, Lin & Ye (2007, 2009, 2012), Lin (2010), Ardakani, et al. (2015), Lucotte (2012)

recent studies but also with other studies from the period. However, following the timeline, we can observe that already in the early 2000s literature is finally able to accommodate greater samples and implement more sophisticated treatment evaluation methods. This should play in favour of better outcome quality and outcome relevancy of the recent studies compared to the early ones (Lin & Ye, 2007).

Ammer & Freeman (1995) survey experience of three industrial countries that were first to adopt IT regime, i.e. New Zealand, Canada, and the United Kingdom. The study finds that although the countries were successful in attaining their inflation goals, their bond yields suggest that long-term inflation expectations for these countries persistently tended to exceed long-term targets. Therefore, it appears that the countries did not achieve full credibility at least in the first few years after IT adoption.

Similar country sample was assessed by Mishkin & Posen (1997) who examine the adoption, operational design, and experience of IT as a framework for monetary policy among New Zealand, Canada, and the United Kingdom. Furthermore, they also analyse the German monetary regime, which according to authors incorporated many of the same features as later IT regimes. The comparative study finds that after IT adoption, all of the assessed countries have maintained low rates of inflation and succeeded in increasing transparency of their monetary policymaking without adversely affecting business cycle development. However, there is no evidence that the IT countries performed better in any aspects than non-IT countries.

Groenevald, et al. (1998) address the issue of whether a switch to IT regime can help achieve lower inflation and increase monetary policy credibility. The study evaluates the success of IT in Canada, New Zealand, and the United Kingdom. To investigate to what extent the joint dynamic processes of inflation and nominal interest rates in these three countries have experienced a structural break at the time of the IT adoption they are paired with the United States, Australia, and Germany. The study finds that there is no clear evidence that IT is superior to other monetary regime and that it performed better in lowering inflation and increasing central bank’s credibility than the other monetary regimes did.

Bernanke, et al. (1999) comprise detailed case studies of the experiences of IT adoption for New Zealand, Canada, the United Kingdom, and Sweden. These are being compared to Switzerland and Germany which are historically known to be money growth targeters but the authors classify them as pseudo inflation targeters. The study finds evidence that IT countries were successful in lowering actual and expected inflation however were unsuccessful in decreasing sacrifice ratios. There is however no evidence that the IT countries performed any better than non-IT countries.

Cecheti & Ehrmann (2002) ask whether or not IT adoption increases a country's output-volatility. They estimate the change in the preferences of central banks in order to see whether the outcomes in IT countries are different compared to non-IT countries due to change of policymaker's objective function. The hypothesis is that aggregate shocks create a trade-off between output and inflation variability. The dataset includes 23 countries out of which 9 are IT countries. To estimate the structural responses for each country to a monetary policy shock, the study implements structural vector auto-regressions. Results suggest that aversion to inflation decreased during the 1990s in all assessed countries, regardless of whether or not they adopted IT. However, the revealed aversion appears to increase more for IT countries.

Pétursson (2004) analyse a sample of 21 IT countries, including developing countries. He evaluates the effect of IT adoption on a set of macroeconomic outcomes using univariate AR(2) model and a dummy variable for pre and post-IT periods on a country basis. He finds that IT adoption has been beneficial to reduction of the inflation level, inflation persistence, and inflation variability. Furthermore, these results have not been at the cost of lower growth or increased variability of business cycle. However, the methodology provided by this study does not tackle the fundamental question of relative difference in performance between IT and non-IT countries. It does not provide a clear and robust evidence of the comparative benefits of IT.

Levin, et al. (2004) study inflation persistence using five industrial IT countries that are compared to seven industrial non-IT countries. The study performs univariate regressions on inflation for each country and finds that IT adoption plays an important role in anchoring long-run inflation expectations and reduces inflation persistence. On the other hand, Levin

& Piger (2004) under similar empirical methodology with sample of twelve industrial countries allow for structural breaks in the univariate regression and find that IT does not seem to have a large impact on long-term inflation expectations.

Kuttner & Posen (1999) conducts time series analysis of inflation properties before and after IT adoption for the United Kingdom, Canada, and New Zealand. In all cases, the experience under an IT regime is compared with the period before its adoption. Study shows that IT adoption in all assessed countries was associated with a reduction in both the level of inflation and inflation persistence without an increase in the relative weight on inflation. For the United Kingdom and Canada, lower inflation levels and persistence after IT adoption are combined with greater accommodation of real shocks and more stable private-sector inflation expectations. The results for New Zealand after IT adoption mix reduced inflation level and persistence with less stable inflation expectations.

Johnson (2002) constructs panel and compares five industrial IT countries (Australia, Canada, New Zealand, Sweden, and the United Kingdom) to six industrial non-IT countries (France, Germany, Italy, the Netherlands, Japan, and the United States) to assess the change in behaviour of inflation expectations related to IT adoption. Controlling for country effects, year effects and the business cycle, Johnson finds that the level of expected inflation in targeting countries significantly reduced after IT adoption announcement. However, Johnson also documents that IT has not reduced absolute average forecast errors in IT countries relative to those in non- IT countries.

Neumann & von Hagen (2002) conduct time series analysis (data for 1978-2001), examining the changes of short-term and long-term interest rates and of inflation and output gaps in response to IT adoption for a rather small sample of 9 countries, among which 6 are explicit IT and the remaining 3 represents control non-IT countries (USA, Germany, Switzerland). The results suggest that the IT countries have succeeded in their determination to stabilize inflation over the medium run and to gain credibility. However, there is no evidence that the IT countries performed better than the rest. Following that, the authors perform an event study to quantify the response of inflation and long-run and short-run interest rates to supply shocks (oil shocks in 1978–79 and in 1998–1996). They find that the effect of IT is not significantly different from zero for average inflation, but it is for interest rates. This might signal a gain in credibility among IT countries.

Ball & Sheridan (2003) attempts to measure the effect of IT on macroeconomic performance by examining twenty OECD[19] countries out of which 7 adopted IT framework in the 1990s. The remaining 13 countries are used as a control group. Authors conduct difference in differences, comparing the countries in period before IT adoption and after, documenting that their performance in measured outcome variables improved (inflation, inflation variability, output, output volatility, and volatility of interest rates), however so did performance of the other non-targeting countries. This suggests that the improvement is not linked with IT adoption and that the countries experience "regression to the mean", i.e. countries with high and unstable inflation tend to see the pattern of decreasing inflation and inflation variability regardless of whether they adopt IT or not. Authors argue that their methodology should provide unbiased estimates of the IT adoption effect. However, there are several flaws to their approach. First of all, it is important to question the treatment effect estimation methodology. If there is an IT country with a rather poor macroeconomic performance before IT adoption it should be compared with a non-IT country with equally poor initial performance. Otherwise, the treatment effect can be over or underestimated. Second, since they are regressing the change in the mean of the measured outcome variables in two different periods in Inflation Targeting dummies, they are running a regression with only 20 observations. Using such small sample may cause inability to reject a false null hypothesis[20] (Ardakani, et al., 2015). Furthermore, the before and after sample averages that are constructed to calculate the change in measured outcome variables cover different periods for each country (since each of them adopted IT at different point in time). This might have adverse effect on the estimates as well.

Similar methodology to Ball & Sheridan (2003) was accepted by Wu (2004). Wu tries to avoid the problems of previous works of Neumann & von Hagen (2002) and Ball & Sheridan (2003) by applying multi-period differences-in-differences estimation to the quarterly CPI inflation rates for period from 1985 to 2002 for 22 OECD industrial countries. He argues

---

[19] Organization for Economic Cooperation and Development

[20] The null hypothesis on question is that the coefficient of the IT dummy is equal to zero

that the multi-period differences-in-differences is a more suitable approach for situation when the different subjects (countries) underwent treatment (IT adoption) in different time periods. He finds that countries that have adopted IT experienced a decrease in their average inflation rates that is not attributable only to regression to mean. Furthermore, there seems to be no evidence that IT countries experienced a significant increase in the volatility level of their real interest rates after IT adoption.

3.2 PSM Literature

Vega & Winkelried (2005) represent probably the first study on this topic that tried to reflect current trends in program evaluation literature in terms of treatment selection problem. Authors attempt to estimate the average treatment effect of IT adoption over inflation dynamics using a much wider control group than the previous research did. Dataset includes 23 IT countries and 86 non-IT control countries. Many observations are however missing. In order to mimic properties of natural experiment and to find suitable counterfactual to the actual IT countries they introduce kernel-based propensity score matching technique. First they use a logit regression model to estimate the propensity score, using the following outcome-independent variables: investments to GDP, openness ratio[21], share of World GDP, fiscal balance to GDP, CPI inflation rate, inflation volatility[22] and money to GDP[23]. They use the estimated propensity score for matching and evaluation of the treatment effect on inflation, inflation variability and inflation persistency. The study finds that IT adoption delivers the outcomes that are "promised" by theoretical literature, i.e. lower mean inflation and lower inflation volatility. Inflation persistence appears to be lower as well. However, the treatment effect estimates are of small magnitude and in many cases statistically insignificant. It is worth noting that there are several drawbacks to authors' methodology. First of all, densities of the estimated propensity score for IT and non-IT countries are highly different which indicates possible problems with insufficient overlap and with area of common support. This might be particularly exaggerated by implementation of kernel matching which uses all control observations to calculate the treatment effect. Lack of sufficient overlap and inconsistent area of common support may lead to unreliable treatment

---

[21] Sum of imports and exports to GDP

[22] Standard deviation of a 5 year moving average.

[23] M2 and M3 money aggregates (vary based on availability)

effect estimates due to bad matches (Austin, 2011b). Furthermore, Vega & Winkelried do not address balancing properties of the model and we are therefore unable to assess its quality and reliability[24].

Another rigorous attempt to address the self-selection problem has been done by Lin & Ye (2007). Authors estimate the average treatment effects of IT adoption using various propensity score matching methods to mimic randomized experiment – i.e. crucial precondition to treatment effect assessment. The study includes 321 annual observations for seven industrial IT countries and 15 non-targeting industrial countries for the period from 1985 to 1999. The measured outcome variables are inflation, inflation variability, long-term nominal interest rates variability, and the variability of income velocity of money. Firstly, authors conduct propensity score estimation using probit regression model incorporating the following outcome-independent variables: lagged inflation rate, trade openness[25], broad money growth[26], fiscal balance, real per capita GDP growth and a five-year central bank governor turnover rate as an inverse proxy of central bank independence. The study finds that the probability of IT adoption is adversely affected by lagged inflation, broad money growth, trade openness and by low central bank independence. Contrary to that, the real per capita GDP growth appears to have positive effect on this probability. Implemented matching techniques include Nearest neighbor, Radius and Kernel matching. The results of PSM shows that effects of IT adoption on inflation and inflation variability are quantitatively small and statistically insignificant in the assessed seven countries. The treatment effects on long-term nominal interest rates and income velocity of money are found insignificant as well. Authors point out that although no non-IT country from their control dataset publicly announced any inflation target, some of them implement policies that are very similar to those of IT countries. It is worth noting that the study does not incorporate any robustness testing that would assess balancing properties, overlap, or area of common support. We are therefore unable to assess quality and reliability of the models.

[24] For explanation of importance of *balancing properties, overlap* and *area of common support* assumptions see section 7.2

[25] Measured as sum of exports and imports to GDP

[26] M3 money aggregate

Lin & Ye (2009) are following in the previously lined-up methodology, however they steer their focus towards developing countries. The idea is that an announcement of an IT adoption makes a central bank's policy more credible, which should help lower inflation expectations, inflation and inflation variability. Authors assume that credibility of central banks in developing countries is significantly lower than that in industrial countries, and therefore suspect that the credibility gain from explicit announcement of an IT adoption would be much more substantial in developing countries (Lin & Ye, 2009). As in the previous case, the study implements a variety of propensity score matching methods to mimic the properties of natural experiment, including Nearest Neighbor, Radius, Kernel matching and local linear regression. Propensity score estimation is conducted incorporating the following outcome-independent variables: lagged inflation rate, trade openness, broad money growth, and real per capita GDP growth. Similarly to previous study the lagged inflation, trade openness and broad money growth are found to have adverse effect on IT adoption probability while the real per capita GDP growth is found to affect the probability positively. Results imply that on average, IT adoption has large and significant effects on lowering both inflation and inflation variability in the assessed 13 developing countries. Furthermore, authors conduct robustness testing by eliminating data samples for periods of hyperinflation[27] and by implementing alternative IT adoption dates obtained from Rose (2007). The results are found to be similar even after these adjustments. However, it is again important to note that apart from aforementioned the robustness testing does not include checks for balancing properties, nor for overlap or area of common support. We are therefore unable to assess quality and reliability of the models.

Lin (2010) attempts to take the analysis a step further by incorporating data for both industrial and developing countries and assessing both of them jointly as well as separately and by focusing on other outcome-variables than the previous research did[28]. The dataset includes annual observations for 22 industrial countries and 52 developing countries for the period of 1985 to 2005. The data include 10 industrial IT countries and 13 developing IT countries. Lin uses probit regression to estimate the propensity score incorporating the same

[27] Annual inflation rates higher than 40%

[28] Nominal and real exchange rate variability and central bank’s reserve to GDP proportion

outcome-independent variables as Lin & Ye (2009)[29]. The lagged inflation, broad money growth, and trade openness are again found to have adverse effect on probability of IT adoption, while real per capita GDP and fiscal surplus appear to affect it positively. The average treatment effect is once again estimated using the same propensity score matching techniques as in Lin & Ye (2009)[30]. The results show that the overall effects of adopting IT are insignificant in pooled dataset, however the results significantly differ between industrial and developing countries. IT significantly increases real and nominal exchange rate stability and international reserves in developing countries but lowers them in industrial countries. The study does not include robustness testing dedicated to balancing properties, overlap, or area of common support.

Lin & Ye (2012) assess effect of IT adoption on inflation rates focusing only on developing countries. The authors incorporate a large sample of 50 countries[31] for period of 1990 to 2006. Properties of natural experiment are mimicked with PSM techniques. Furthermore, the study implements dynamic panel generalized method of moments (hereinafter "GMM") regressions to provide additional evidence. The propensity score is estimated using probit regression model incorporating outcome-independent variables similar to previous research[32]. Lagged inflation, trade openness and broad money growth are found to have adverse effect on IT adoption probability. On the other hand, size of economy and real per capita GDP growth tend to affect it positively which is in line with the previous research. Using both PSM and GMM the study finds strong evidence that, compared to exchange-rate targeting, IT leads to a significantly lower inflation rate, which does not come at a cost of lower economic growth. Once again the study omits robustness testing dedicated to balancing properties, overlap and area of common support.

---

[29] Lagged inflation, broad money growth, real per capita GDP growth, fiscal balance and trade openness

[30] Including Nearest Neighbor, Radius, Kernel matching and local linear regression

[31] Out of which 13 are IT countries

[32] Trade openness, economy size, lagged inflation, broad money growth, central bank's turnover rate, real per capita GDP growth. Furthermore, the authors implement the following control variables that are included to the benchmark model: Reserve ratio, current account balance, fiscal balance, financial openness and inflation of the United States as a proxy for world inflation

Lucotte (2012) investigates whether IT adoption affects the fiscal effort in emerging markets economies. Study implements PSM techniques to mimic properties of natural experiment. The dataset comprises of 515 observations for 30 countries out of which 14 are IT countries. To estimate the propensity score, study introduce probit regression model and among other include the following outcome-independent variables: real per capita GDP, lagged inflation, trade openness, credit provided by domestic financial institutions. The propensity score estimates indicate that the probability of IT adoption is positively affected by real per capita GDP growth, and negatively by lagged inflation and trade openness. These estimates are consistent with previous research. Estimates for credit provided by domestic financial institutions are mixed and statistically insignificant. Implemented PSM techniques include Nearest neighbor, Radius and Kernel matching. Matching results indicate that, on average, inflation targeting has a significant positive effect on public revenue collection. Furthermore, the study suggest that sound fiscal policy is not a fundamental precondition for the adoption of IT, relative to other prerequisites such as central bank independence or the flexibility of the exchange regime (Lucotte, 2012). Author conducts robustness testing first by introducing three additional variables to the propensity score estimation model that could simultaneously influence the choice of adopting IT, second by omitting time periods from before 1990 and finally by omitting periods of hyperinflation. Nonetheless the study does not include analysis of balancing properties, overlap or area of common support and we are therefore unable to assess quality of the estimates.

Minea & Tapsoba (2014) explore countries‘ performance after IT adoption in terms of their fiscal discipline. The study incorporates datasest of 84 countries out of which 62 are classified as developing and 22 as industrial countries. The dataset includes observation over the period of 1985-2007. Among the 84 countries, 30 are identified as inflation targeters, the rest are controls. The propensity score is estimated using a probit regression and among others the following outcome-independent variables: lagged inflation rate, broad money growth, trade opennes, lagged debt-to-GDP ratio or GDP growth rate. The propensity score estimates imply that higher lagged inflation, higher broad money growth, higher GDP growth, and greater trade opennes have negative effect on probability of IT adoption. The ATT effect on fiscal discipline is estimated using various PSM methods, including Kernel or Stratification matching. The estimates show that IT adoption significantly improves fiscal discipline in IT countries compared to non-IT countries. The study conducts robustness testing by introducing alternative measures for fiscal discipline, alternative specification of the propensity score model, and by introduction of alternative IT adoption dates. ATT estimates remain stable even after these adjustments. However, similarly to the previous research the study does not include any analysis of balancing properties, overlap or area of common support and we are therefore unable to assess quality and reliability of the estimates.

Ardakani, et al. (2015) use annual data of 98 countries for the period from 1990 to 2013. Among the sample there are 27 IT countries and 71 non-IT countries. Authors use a logit regression model to estimate the propensity score incorporating the following outcome-independent variables: real per capita GDP growth, money growth[33], lagged inflation, trade openness[34], financial depth[35], and central bank's assets to GDP. Estimates are somewhat inconsistent with the previous research done by Lin & Ye (2007, 2009, 2012) and Lin (2010). Study implies that the probability of IT adoption is negatively affected by real per capita GDP growth for both industrial and developing countries and positively affected by money growth in both developing and industrial countries. Furthermore, the effect of lagged inflation appears to be negative in industrial countries, however positive in developing countries and in both cases statistically insignificant. Study argues that models used in the previous research might have suffer from misspecification. To control for this, authors introduce non-parametric and semi-parametric single index matching. The results indicate that IT lowers inflation, inflation variability and improves fiscal discipline in both developing and industrial countries. IT regime appears to negatively affect interest rates volatility. However, balancing properties of the model are disputable. The matching did not succeed in reducing the mean standardise bias in all observed covariates between treatment and control group[36]. The mean bias in observed covariates between treatment and control group are in most cases greater than 10 per cent and for one variable the matching did not even succeed in reducing the bias. This might mean breach of conditional independence assumption which requires that the potential outcomes are independent of treatment

[33] M3 monetary aggregate

[34] Sum of exports and imports to GDP

[35] Proxy for level of development of domestic financial sector

[36] I.e. the difference in covariate means between control and treatment group were lower prior to matching. This signals that matching might not have been successful (Caliendo & Kopeinig, 2005).

assignment, i.e. that treatment selection is solely based on observable characteristics and that all variables that influence treatment assignment and outcome simultaneously are observed (Caliendo & Kopeinig, 2005). Furthermore, the study does not omit extreme observations such as periods of hyperinflation or of abnormal GDP growth from the sample. This might have adverse effects on the estimates.

4 Dataset

In this section I present the datasets and the variables used for estimation of propensity score and of ATT for the respective outcome variables.

The original dataset was constructed as an unbalanced panel[37]. To be able to conduct PSM, I eliminate the time dimension from the data by transforming it into long-form quasi cross-sectional sample. To ensure proper estimation of propensity score, the distribution of observed covariates among matched sample has to be similar between treated and control groups[38] – i.e. *conditional independence assumption* (see part 5.2 for detail). Significant differences in the covariates' distribution after the matching might be a signal of potentially biased estimates[39] (Austin, 2011b). In order to abide these balancing properties I conduct further adjustments to the obtained sample. A few countries in the dataset have experienced high rates of inflation (higher than annual rate of 30 per cent). Such observations were eliminated from the sample to avoid estimate bias[40]. Similarly, a few countries experienced a drastic year by year decrease of GDP levels[41] that usually lead to its abnormal growth rates in the following periods. Such observations were eliminated from the sample as well. Furthermore, I also eliminated several countries that might be referred to as "extreme" observations[42].

The final pooled dataset includes annual observations of 78 countries (1,968 observations) for the period from 1984 to 2015. Majority of the data is obtained from the World Bank's World Development Indicators (hereinafter "WDI")[43]. Among the sample there are in total 25 IT countries[44] (*treated*) and 53 non-IT countries (*control group*).

---

[37] Some observations for output-independent variables are missing. See Table 13 – 21 (appendix) for descriptive statistics.

[38] More on balancing properties of the model can be found in section 577.2

[39] This might be for example the case for Ardakani, et al. (2015)

[40] Same applies for observations with lagged inflation rate higher than 30%.

[41] Typically associated with a military conflict (for example Egypt and Libya).

[42] Eliminated countries include Liechtenstein, Luxembourg, Malta, Singapore, and Puerto Rico. I.e. apart from Puerto Rico very small industrial countries with high levels of real per capita GDP.

[43] Broad money growth for all Eurozone countries was obtained from Eurostat.

[44] See section 4.3 for details.

According to Aizenman, et al. (2011) there are many reasons why it is important to distinguish between industrial and developing countries when assessing the effect of IT adoption[45]. Both groups may have different institutional arrangements, may highly differ in institutional credibility and political independence of the central bank, and also may have different inflation and macroeconomic histories. Not reflecting this may lead to biased estimates (Aizenman, et al., 2011).

Table 1 - Summary of Developing countries

Developing countries			
Non-targeting countries		Inflation targeting countries	IT Adoption
Albania	Lebanon	Armenia	2006
Argentina	Liberia	Brazil	1999
Azerbaijan	Libya	Chile	1999
Bahrain	Lithuania	Colombia	1999
Bangladesh	Malaysia	Czech Republic	1998
Belarus	Moldova	Hungary	2001
Bolivia	Montenegro	Indonesia	2005
Bosnia and Herzegovina	Paraguay	Israel	1992
Bulgaria	Puerto Rico	Korea, Rep.	1998
Costa Rica	Qatar	Mexico	2001
Croatia	Saudi Arabia	Peru	2002
Egypt, Arab Rep.	Slovak Republic	Poland	1998
Estonia	Slovenia	Romania	2005
Georgia	Tunisia	Serbia	2006
China	Ukraine	South Africa	2000
India	United Arab Emirates	Thailand	2000
Jamaica	Uruguay	Turkey	2006
Kazakhstan	Vietnam		
Latvia			
n = 36		*n = 17*	

The pooled dataset is therefore further divided into two separate datasets, one for developing countries and the other one for industrial countries. Each country is allocated to either group based on its level of economic development[46]. This allows me to assess the effect of IT adoption separately for both groups and avoid potential estimate bias.

The list of developing countries is presented in Table 1, industrial countries are listed in Table 2. Both tables also contain period in which IT was first introduced in the given IT country. There are in total 17 inflation targeters and 36 non-targeters among developing

[45] Further discussed also in section 2.2

[46] See section 4.2 for details.

countries and 8 inflation targeters and 17 non-targeters among industrial countries. Later, in section 8, I conduct robustness testing by altering this assignment into industrial/developing group as well as a country's identification as an IT/non-IT.

Table 2 - Summary of Industrial countries

Industrial countries			
Non-targeting countries		Inflation targeting countries	IT Adoption
Austria	Italy	Australia	1993
Belgium	Japan	Canada	1991
Cyprus	Netherlands	Iceland	2001
Denmark	Estonia	New Zealand	1990
Finland	Portugal	Norway	2001
France	Singapore	Sweden	1993
Germany	Spain	Switzerland	2000
Greece	United States	United Kingdom	1992
Ireland			
n = 17		*n = 8*	

4.1 Variables

Constructing propensity score matching model requires dedication of a treatment variable, of an outcome-dependent variable(s) and of several outcome-independent variables that are going to be used for the estimation of propensity score.

The dependent variable for the estimation of propensity score is the dummy variable for inflation targeting. The variable equals 1 for countries that are fully-pledged explicit inflation targeters. I estimate the ATT for inflation, inflation variability, annual real per capita GDP growth and its variability (outcome-dependent variables). Following Lin & Ye (2007) and Ardakani, et al. (2015), I measure inflation variability and real GDP per capita variability by the standard deviation of its three-year moving average.

Outcome-independent variables used for the estimation of propensity scores are broad money growth, current account balance, lagged inflation rate, economy size, financial depth, fiscal balance, net national income per capita, population growth, reserve ratio, and finally trade openness. Summary of all variables can be found in Table 3, outcome-independent variables are discussed in detail in section 5.4.

Table 3 - Summary of variables

Variable	Type	Description	Source
IT	Treatment	Inflation targeting – dummy variable	--
RGDPpcGr	Dependent	Annual % growth of real Gross Domestic Product per capita	WDI
RGDPpcVAR	Dependent	Standard deviation of a three-year moving average of *RGDPpcGr*	WDI
log_CPI	Dependent	Common logarithm of annual inflation measured with CPI	WDI
CP_infl_VAR	Dependent	Standard deviation of a three-year moving average of *log_CPI*	WDI
BMGr	Independent	Broad money growth (M3, annual %)	WDI & Eurostat
BoP	Independent	Current account balance as a % of GDP	WDI
CPI_infl_1	Independent	*log_CPI* lagged by 1 period	WDI
Esize	Independent	Ratio of country's GDP to the World GDP	WDI
FinD	Independent	Domestic credit provided to private sector as a % of GDP	WDI
FisBal	Independent	Cash surplus/deficit as a % of GDP	WDI
NNIpcGr	Independent	*Adjusted net national income per capita growth*[47]	WDI & IFS[48]
PopGr	Independent	Annual % population growth	WDI
ResR	Independent	Ratio of total reserves minus gold to GDP	WDI
Trade	Independent	Sum of exports and imports as a % of GDP	WDI

4.2 Industrial and developing countries

Before assessing the effect of IT adoption, it is important to distinguish between developing and industrial countries. It is well documented that they might significantly differ in their institutional arrangements, central bank independency, and other aspects. IT adoption is therefore unlikely to have the same effect on both groups (Aizenman, et al., 2011). Lin & Ye (2009) assume that an announcement of IT adoption makes a central bank's policy more credible, which should help lower expected inflation, inflation, and inflation variability. Since the credibility of a central bank in developing country is supposed to be lower than that in industrial country the authors expect that the credibility gain from an explicit announcement of IT adoption should be more substantial for developing countries. Calvo & Mishkin (2003) specified six fundamental institutional deficiencies that emerging countries usually face. These include: weak financial institutions, weak government prudential regulation and supervision, low credibility of monetary institutions, currency substitution,

[47] Adjusted Net Nation Income is calculated as Consumption + Investments + Govt. Spending + Net Export + Net Foreign Factor Income - Indirect Taxes - Manufactured Capital Depreciation - Natural Resource Depletion

[48] International Monetary Fund's International Financial Statistics database

tendency for liability dollarization, and vulnerability to sudden economic stops. While industrial countries are not immune to the aforementioned problems, the relevant magnitude of the problems is much lower[49].

In order to accommodate this into my research, each country is assigned to either developing or industrial dataset based on its level of economic development.

There are no commonly agreed parameters based on which a country can be classified as either emerging or industrial. Furthermore, there is a variety of sources that provide different assignments for emerging and industrial countries which are sometimes even contradictory to each other[50]. I conduct the assignment based on three independent sources: a generally well-known J.P. Morgan's Emerging Bond Index (hereinafter "EMBI") (J.P. Morgan, 2016), FTSE's Annual Country Classification Review (FTSE, 2016) and International Monetary Fund's World Economy Outlook (IMF, 2016).

The EMBI is probably the most favourite benchmark index for measuring the total return performance of international government sovereign bonds issued by emerging market countries. Its portfolio is being regularly updated to reflect contemporaneous economic development (J.P. Morgan, 2016). However, EMBI's shortcoming is that its portfolio is somewhat limited and does not cover all emerging markets included in my dataset. I use the FTSE's Annual Country Classification Review to gain some overlap for countries that are missing in EMBI's portfolio, however the FTSE's list does not cover all of my developing countries either. That is why I further incorporate methodology of IMF's World Economy Outlook. Combination of all three sources provides me with a sufficient overlap and allows me to assign countries that are generally accepted as developing but are missing in the FTSE's review or in EMBI's bond portfolio. There are several discrepancies between the sources. While FTSE classifies the Czech Republic, Israel, and Slovenia as emerging countries, EMBI and IMF accounts all three of them as Industrial. It is the other way around for Greece which is considered emerging by EMBI but industrial by FTSE and IMF. Another Example is the Republic of Korea, which is considered developing by EMBI but industrial by IMF and FTSE. The complete list of countries classified as emerging by each of the sources is presented in Table 11 in the appendix. To check for possible

[49] For detailed discussion on differences between emerging and industrial countries see section 2.2

[50] See detail of contradictory assignments in Table 11.

misclassification, I conduct robustness tests, including reclassification of all aforementioned discrepancies in section 8.1.

4.3 Targeting Criteria

A major problem with which the contemporaneous literature still struggles are the multi various characteristics of IT. Miao (2009) argues that a closer look reveals that divergence among IT regimes is rather norm than an exception even for the more homogenous group of so-called fully-fledged inflation targeters. There are also other very important differences among central banks such as reaction horizon, transparency and accountability (Miao, 2009). Furthermore, it appears that there is not a broad common ground on what constitutes IT. Kuttner (2004) suggested that there are two main ways to think about IT: first, in terms of the observed characteristics of the policy framework and second in the terms of policy rule. While Svensson (1996, 1999) emphasizes the latter, Bernanke, et al. (1999) suggests that IT is better described as a framework rather than a rule.

There are similar discrepancies in the terms of IT components. Mishkin (2004) lists five core components to an inflation targeting regime: I. The public announcement of an explicit inflation target, II. An institutional commitment to price stability as the primary goal of monetary policy, III. An information-inclusive strategy to set policy instruments, IV. High transparency of the monetary policy, and V. Increased accountability of central bank for attaining its inflation objectives. However, Truman (2003) provides shortened list of four: I. Price stability as the primary goal of monetary policy, II. Publicly announced explicit numerical inflation target, III. A time horizon over which the target is to be met, IV. An associated approach for objective evaluation.

Notwithstanding the aforementioned discrepancies, recent literature appears to find common ground at least on the definition of a lowest common denominator for recognition of IT. IT must have a well-defined (numerical) inflation target with institutional arrangements to support its achievement, and a high degree of transparency and credibility. It must also establish and maintain well-anchored inflation expectations (Miao, 2009).

It will probably come as no surprise that empirical research often cannot consent on a country's identification as either IT or non-IT. For example, Carare & Stone (2006) and

Clardia, et al. (1998) identify the G-3[51] countries as implicit targeters[52] and argue that the low inflation levels they achieve might be attributable to their IT-like monetary policies. Truman (2003) on the other hand defines G-3 as a whole new separate group of countries that are not inflation targeters yet they are able to achieve well-anchored inflation expectations just as IT countries do. Even if we find a consensus on central bank's identification as inflation targeter or non-targeter another issue arises with the dating of the IT adoption.

Most recently the identification and dating problem were addressed by Miao (2009) and Aizenman, et al. (2011). Miao identifies 21 de jure explicit inflation targeting countries and conducts their comparison in terms of flexibility and transparency. Aizenman, et al. identifies fully-fledged inflation targeters among developing countries and assesses the role of the real exchange rate on their macroeconomic performance. Furthermore, Hammond (2012) provides us with an exhaustive handbook on 27 inflation targeters from both industrial and developing countries and provides detailed description of their policy framework.

Following the recent works of Lin (2010), Lin & Ye (2012), and Ardakani, et al. (2015) I obtain a country's identification and its IT adoption date from Rose (2007), Miao (2009) and Hamond (2012). In compliance with the methodology of Lin (2010), Lin & Ye (2012), and Ardakani, et al. (2015) I classify all member countries of the Eurozone as non-IT due to lack of explicit IT related announcement from the European Central Bank. This also includes observation for countries that in the past had their own monetary policy, were at some point IT, but afterwards joined the Eurozone[53]. The full list of IT countries (including the adoption dates) from all three sources is presented in Table 12 in the appendix. Discrepancies between the sources are addressed and tested in section 8.1.

[51] USA, EMU, Japan

[52] i.e. the countries implement several identical policies as IT countries, however without public announcement of the explicit numerical inflation target.

[53] I.e. Spain (1995-1999), Finland (1993-1999), and Slovakia (2005-2008)

5 Methodology

Recently, observational studies have been frequently used for treatment effect estimation. In randomized controlled experiment, randomization ensures that, on average, treated subjects do not systematically differ from control subjects in both measured and unmeasured characteristics. The treatment effect can be therefore estimated directly by comparing outcomes of both groups. However, non-randomized experiments can be subject to selection bias in which treated subjects systematically differ from control group. The treatment effect is therefore not as straightforward to estimate (Austin & Stuart, 2015).
The central idea of propensity score matching is to mimic the properties of randomize experiment. The approach works as a substitute to systematic methods of experimental design in cases where no such approach for constructing control group could have been applied. In this part I further discuss reasons for introduction of this method rather than of common linear regression for estimation of the causal effect. Furthermore, I describe my approach towards propensity score estimation and assumptions that must apply. In the end I conduct the estimation of propensity scores itself.

5.1 Propensity Score Matching

It is well established that estimation of causal effect that is obtained by comparing treatment and control group where no systematic method of experimental design is used to construct the control group might be biased due to problems such as self-selection or endogeneity of treatment[54]. If decision to target inflation is not random and it is in fact systematically correlated with a set of observables that also affect the outcomes, then we face the *selection on observable* problem which renders common linear regression unreliable (Lin & Ye, 2007; Dehejia & Wahba, 2002; Heckman, Ichimura, & Todd, 1997). In the last two decades, PSM methods have been widely used in statistics literature to overcome such problems[55]. However, the methods appeared to be unknown to economics literature only until recently (Dehejia & Wahba, 2002).

[54] Mishkin & Schmidt-Hebbel (2001) and Gertler (2005) point out that IT adoption is clearly an endogenous choice.

[55] See for example: Cave & Bos (1995), Czajka, et al. (1992), Rosenbaum & Rubin (1985) or Rosenbaum (1995).

Matching is a rather straightforward approach that can be used in order to eliminate problems of non-experimental nature, such as selection bias and to effectively mimic properties of natural experiment (Cochran & Chambers, 1965).

It is a method under which the treated subjects are matched with control subjects for further assessment in such a way so that they are both similar in terms of their observable characteristics (Rubin, 1973). When the relevant differences between two subjects are captured in the observable covariates, matching methods can provide us with unbiased estimate of the treatment effect (Dehejia & Wahba, 2002).

Applying this logic to our case, we would simply match an inflation targeting subject with a control subject, judged on their similarity in observed covariate X (for our purposes financial depth or some other variable). In practice, however, conditioning on all relevant covariates, as their number increases, is limited by its high-dimensionality[56] and therefore it might not be feasible to apply this method. That is why it is appealing and also crucial to apply the propensity score matching, under which the balance of high-dimensional covariates can be achieved without requiring every single covariate to be "similar". Rosenbaum and Rubin (1983) establish that if matching on covariate X is valid (i.e. potential outcomes are independent of treatment conditional on covariates X[57]) so is valid matching solely on the balancing score X, which for our purposes is the probability of selection into treatment, *i.e. propensity score*[58] (Rosenbaum & Rubin, 1983, p. 44):

$$P_r(D = 1|X) = P(X) \tag{5.1}$$

This allows us to simplify complicated multi-dimensional matching into a one-dimensional problem (Rosenbaum & Rubin, 1983).

5.2 Treatment Effect & Key Assumptions

Propensity score matching allows us to estimate the ATT (Imbens, 2004). If the outcome is continuous, the treatment effect can be estimated as the difference between the mean

---

[56] i.e. if X contains n covariates, the number of possible matches will be 2^n

[57] Assumption further addressed in section 5.2

[58] Propensity score is defined as the probability of assignment into treatment conditional on measured covariates. It is balancing when the distribution of propensity score among treated and control group is similar (Rosenbaum & Rubin, 1983).

outcome for treated subjects and the mean outcome for control subjects in the matched sample (Rosenbaum & Rubin, 1983). In order to evaluate the treatment effect of inflation targeting in targeting countries we denote (Caliendo & Kopeinig, 2005, p. 3)

$$ATT = E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 1] \tag{5.2}$$

where D is the targeting dummy, $(Y_{i0}|D_i = 1)$ is the counterfactual value of the outcome that would have been observed had not the country adopted IT policy, and $(Y_{i1}|D_i = 1)$ is the outcome observed in the same country. An obvious problem with this estimation is that the right-hand side of the equation remains unobserved. We cannot observe country's inflation variability or per capita GDP variability had the country not adopted the policy; however, it can be estimated. In order to do so, the *conditional independence assumption* has to apply (Rosenbaum & Rubin, 1983, p. 42):

$$(Y_0, Y_1 \perp D|X) \tag{5.3}$$

It requires that, conditional on *X*, the potential outcomes are independent of treatment assignment. This implies that treatment selection is solely based on observable characteristics and that all variables that influence treatment assignment and outcome simultaneously are observed. Taking this into account, equation 5.2 can be rewritten as follows (Caliendo & Kopeinig, 2005, p. 4):

$$ATT = E[Y_{i1}|D_i = 1, X_i] - E[Y_{i0}|D_i = 0, X_i] \tag{5.4}$$

Furthermore, Heckman, et al. (1998) establishes that for fundamental identification of the mean treatment effect (*ATT*) just the *weak mean independence condition* suffices. This can be denoted as follows (Heckman, et al., 1998, p. 263):

$$E[Y_{i0}|D_i = 1, X_i] = \mathrm{E}[Y_{i0}|D_i = 0, X_i] \tag{5.5}$$

Here the distribution of Y_0 given X_i for treated subjects can be identified with using only observations of control subjects provided that $X_i \in S$, where *S* is the area of common support. Furthermore, under these conditions it is not necessary to make assumptions about specific functional forms of outcome equations or distributions of the unobservables (Heckman, et al., 1997).

Another additional assumption has to apply to allow for propensity score matching. Under this assumption, every subject from the population with the same X values have a positive probability of being both treated and non-treated[59]. This assumption rules out the phenomenon of perfect predictability of D given X[60] (Caliendo & Kopeinig, 2005, p. 4; Rosenbaum & Rubin, 1983, p. 45):

$$P(D = 1|X) < 1 \tag{5.6}$$

When equations 5.3 & 5.6 hold, we can say that the treatment assignment is strongly ignorable (Rosenbaum & Rubin, 1983). After that the ATT can be estimated as follows (Caliendo & Kopeinig, 2005, p. 4):

$$ATT = E[Y_{i1}|D_i = 1, p(X_i)] - E[Y_{i0}|D_i = 0, p(X_i)] \tag{5.7}$$

5.3 Probit Regression

Propensity score matching is a 2 step process. First we need to estimate the propensity score that will work as a balancing score and will allow us to conduct the actual matching. The propensity score is usually estimated using logit or probit regression models[61]. Both methods are a type of generalised linear models where the dependant variable can only take two values (for our purposes IT country or non-IT country). In both cases the goal is to estimate the probability of a binary response (for our purposes the probability of IT adoption) for an observation with particular characteristics specified by one or more independent covariates (so called "predictors"). The main difference between logit and probit models is in their link function. While Probit model usually takes form of equation 5.9, Logits are generally characterized as (Walker & Duncan, 1967, p. 169):

[59] Caliendo & Kopeinig (2005) and Li (2012) call this *common support* or *overlap assumption*

[60] For estimation of Average treatment effect on the whole population the condition is: $0 < P(D = 1|X) < 1$ (Rosenbaum & Rubin, 1983, p. 45):

[61] Also known as logistic and probabilistic regression.

$$P(Y_{it} = 1|X_{it}) = \left[1 + e^{-X'\beta}\right]^{-1} + \eta_{it} \tag{5.8}$$

There is no general consensus over which model to prefer. It is well documented that both methods, notwithstanding the difference in their link function, perform equally well in binary treatment probability estimation and they tend to yield highly similar results (Hahn & Soyer, 2005). Hence, the choice is not critical for quality of the propensity score estimation (Caliendo & Kopeinig, 2005).

I use the following Probit regression model to estimate the propensity score, which is afterwards used for the actual matching and estimation of ATT:

$$P(Y_{it} = 1|X_{it}) = \Phi(X'_{it}\beta) + \eta_{it} \tag{5.9}$$

Where Y_{it} is a dummy variable for inflation targeting, X_{it} is the set of control variables[62], Φ is the cumulative function of the standard normal distribution, and η_{it} is the error term.

5.4 Propensity Score Estimation

In this section I conduct the first stage estimation of propensity score and examine the role of macroeconomic performance on the likelihood of IT adoption.

Since the propensity scores are unknown, an important first step from the assignment-based perspective is to estimate them. The goal is to obtain estimates of propensity score for all three datasets that statistically balances the covariates between treated and control groups. For that purpose, I use probit regression introduced in the previous section. In all cases the dependant variable is the dummy variable for IT. I introduce several outcome-independent variables as control variables, most of which were already proposed for this purpose in the previous research[63]. I use the variables described below.

In previous research, economic growth is probably the most commonly used macroeconomic characteristic for propensity score estimation. Notwithstanding that there appears to be no clear evidence on its effect on IT adoption probability. Lin (2010) and Lin & Ye (2007, 2012) find positive relation between real per capita GDP growth and the probability of IT

[62] The complete list of control variables is discussed in section 5.4

[63] For example, Lin & Ye (2007, 2009, 2012), Lin (2010) or Ardakani, et al. (2015)

adoption, which is however statistically insignificant. On the other hand, the estimates for real per capita GDP growth in Ardakani, et al. (2015) are negative and statistically significant.

Because I already use real per capita GDP growth as outcome-dependent variable, I introduce adjusted net national income per capita growth instead (hereinafter "*NNIpcGR"*). *NNIpcGR* allows us to assess economic progress while providing a broader measure of national income that also accounts for the depletion of natural resources (Hamilton & Ley, 2010). It is calculated by subtracting from gross national income a charge for the consumption of fixed capital and for the depletion of natural resources. The deduction for the depletion of natural resources reflects the decline in asset values associated with the extraction of natural resources. This is analogous to depreciation of fixed assets. *NNIpcGR* is particularly useful in monitoring low-income, resource-rich economies, because such economies often see large natural resources depletion as well as substantial exports of resource rents to foreign countries (Hamilton & Ley, 2010). Based on previous research of Lin (2010) and Lin & Ye (2007, 2009, 2012) and since there is a certain level of similarity between real per capita GDP growth and *NNIpcGR*, I expect *NNIpcGR* to have positive effect on the IT adoption probability.

Ardakani, et al. (2015) remarks that the previous literature has broadly ignored the role of financial market development on probability of IT adoption. In contrary to that, the theoretical literature on the topic strongly suggests that level of financial market development is one of the most important decision-making criteria. The banking system and capital markets should be sound and well developed to allow for an effective monetary policy transmission (Ardakani, et al., 2015). To incorporate this into the propensity score estimation I make use of variable for Financial depth (*FinD*), which is calculated as the domestic credit provided to the private sector as a percentage of GDP. The variable serves as a proxy measure for degree of domestic financial sector's development. Based on the theoretical literature and the previous work of Ardakani, et al. (2015), I expect positive relation between the level of financial development and the likelihood of IT adoption.

One of the most important elements that decides whether or not an IT regime can work successfully is the central bank's credibility. It is likely that central bank in country with steady inflation level will be more credible and hence more able to create inflation expectations than central bank in country that suffers from high level and high variability of inflation (Lin & Ye, 2007). To take this into account I include the lagged inflation rate

(*CPI_infl_1*)[64] into the model. I expect that countries with higher lagged inflation rates will be less likely to adopt IT regime[65].

Another variable that I incorporate into the model is the trade openness *(Trade)* which is calculated as the ratio of sum of imports and exports to GDP. The effect of trade openness on probability of IT adoption might be less obvious. However, openness may increase benefits of other monetary regimes such as exchange rate targeting on trade (Lin & Ye, 2012). Similarly to that, Romer (1993) suggests that more open economies are less likely to adopt IT. Taking this into account I expect negative relation between trade openness and probability of IT adoption.

Finally, Ardakani, et al. (2015) suggests that other non-macroeconomic variables might also play a vital role in the probability of IT adoption[66]. Following the proposed methodology, I include variable for broad money growth (*BMGr*) as a measure of financial system's health. The broad money growth represents annual growth rate of M3 monetary aggregate. I expect countries with lower broad money growth to be more likely to adopt IT.

It is important to keep in mind that the main goal of estimating propensity score is not to find the best model as possible for explanation of IT adoption probability[67] but to balance all included covariates such that propensity score can serve as a balancing score (Augurzky & Schmidt, 2000)[68]. More important factor is that the estimated propensity scores are distributed widely enough in such way that allows subjects with the same or similar characteristics of propensity score to be observed in both groups[69] (Heckman, et al., 1997). Including additional variables into the model for the sake of better probability explanation

---

[64] Inflation rate lagged by 1 period

[65] Following Mishkin & Schmidt-Hebbel (2001) and Lin & Ye (2012), I expect that other monetary regime such as exchange rate targeting to be more attractive tool for anchoring high inflation for central banks with low credibility.

[66] Specifically, institutional independence, development of technical infrastructure, health of financial system and economic structure (Ardakani, et al., 2015).

[67] Too "good" data, with $P(X) = 1$ (or too bad with $P(X) = 0$) will not allow for matching conditional on those X because all subjects with such characteristics always (or never respectively) receive treatment and matches cannot be performed (Caliendo & Kopeinig, 2005).

[68] Balancing properties are assessed in section 7.2

[69] Overlap testing is conducted in section 7.2.3

may therefore cause an over-parametrisation. This might lead to exacerbation of common support problem[70], increase variance of the estimation and cause its bias (Bryson, et al., 2002). On the other hand, omitting an important variable might also result in biased estimates (Heckman, et al., 1997). In relation to that I introduce another group of control variables in order to conduct robustness testing of the propensity score estimation and to assess whether the benchmark model does not omit an important variable. Each of the additional variables is included to the benchmark model to check for stability of the estimates. The additional variables are: the balance of payments *(BoP),* the size of an economy *(ESize),* the fiscal balance *(FisBal), the* reserve ratio *(ResR),* and finally the annual growth rate of population *(PopGr).*

The balance of payments represents the current account balance as a percentage of the GDP, size of the economy is calculated as the ratio of a country's current GDP to current GDP of the whole world, fiscal balance is the ratio of cash surplus or cash deficit to GDP, reserve ratio is calculated as the ratio of total reserves minus reserves in gold to a country's GDP and finally the population growth is given as an annual growth rate of a country's population.

Lin & Ye (2012) finds that there is a negative relation between *BoP* and the probability of IT adoption (statistically insignificant) for developing countries. There is however no evidence on the effect for industrial countries[71].

According to the previous research the *ResR* and *ESize* also appear to affect the IT adoption probability in a negative way[72] (Lin, 2010; Lin & Ye, 2012).

Mishkin (2004) argues that fiscal stability is a fundamentally necessary condition for inflation control and hence pre-requisite for IT adoption. Irresponsible fiscal policy puts pressure on the central bank to monetize the debt, thereby producing rapid money growth and high inflation (Woodford, 1995). Fiscal imbalances can also lead to banking and financial crises that will not allow any monetary regime to control inflation (Mishkin, 2004). I therefore expect positive relation between fiscal balance and IT adoption probability.

[70] Area of common support is assessed in section 7.2.2

[71] Furthermore, Lin (2010) finds mixed evidence on effect of IT adoption on a country's balance of payments. For industrial countries the relationship appears to be negative but positive for developing countries. In both cases however the estimates are statistically insignificant.

[72] *ESize* is found to have negative but statistically insignificant effect on the IT adoption probability also in Lucotte (2012)

The estimates are reported in Table 4 for pooled dataset, in Table 5 for developing countries and in Table 6 for industrial countries. Most of estimated coefficients have the expected effect on IT adoption probability. Furthermore, addition of another control variable to the benchmark model did not cause any severe variations of the estimates and most importantly in all cases properties of the estimates remained the same. Based on that we can assume that the benchmark model is neither over-parametrised nor omits an important variable.

I find that for all three datasets the lagged inflation rate, the financial depth, and the trade openness systematically and significantly affects a country's probability of IT adoption. The relation appears to be most significant for developing countries. Among developing countries, the significance applies also for net national income and broad money growth.

Positive properties of estimated coefficients for financial depth shows that countries with more developed financial sectors are more likely to adopt IT. This result is consistent with Lucotte (2012) and Ardakani, et al. (2015).

The difference between industrial and developing countries is clearly illustrated by estimates for adjusted net national income and broad money growth. Coefficients of both variables are negative for developing and for pooled datasets. However, the properties are different for industrial countries[73].

Lagged inflation appears to play a crucial role on IT adoption probability. Countries with higher past inflation are less likely to adopt IT. The estimates are highly significant for all three datasets and are consistent with Lin & Ye (2007, 2012), Lin (2010), Lucotte (2012), and Ardakani, et al. (2015).

Estimated relation for Trade openness is in all three cases negative and highly statistically significant. Other monetary regimes such as exchange rate targeting might appear more attractive to countries with high levels of trade openness. The estimates are consistent with Lin (2010) but inconsistent with Ardakani, et al. (2015).

Adding the rest of controlling variables to the benchmark model did not alter properties of any estimates. Negative balance of payments has adverse effect on IT adoption probability for pooled and developing dataset. However, the properties are different for industrial countries. None of the estimates are statistically significant.

[73] Similar results for the real per capita GDP growth are provided by Lin & Ye (2007, 2012) and Ardakani, et al. (2015).

Economy size affects a country's probability of IT adoption negatively in all three cases. This is in line with theory that other monetary regimes might be more attractive for large economies (Mishkin, 2004). The estimates are significant for pooled and industrial datasets. This is consistent with Lin & Ye (2012).

Fiscal balance appears to have positive effect on probability of IT adoption for industrial and pooled dataset. Relation for developing countries is negative[74]. Estimates are in all three cases statistically significant. The results are consistent with Lin & Ye (2012) and for pooled dataset also with Lucotte (2012).

Higher reserve ratio is positively influencing probability of IT adoption in all three datasets. The estimates are in all cases statistically significant. The results are inconsistent with Lin & Ye (2012).

Estimates for population growth are statistically significant, but highly variable.

The overall fit of the performed regressions appears to be reasonable with pseudo R-squared around 0.2. That is comparable to an ordinary least squares' adjusted R-squares of 0.7 (Louviere, et al., 2000).

[74] This might support the view of Mishkin (2004) and Bernanke, et al. (1999), that good fiscal performance does not have to be seen as prerequisite to IT adoption and on the contrary that IT framework can be partly used as a tool for gradual promotion of fiscal and financial reforms (see section 2.2.1 for detailed discussion).

Table 4 – Estimated Propensity Scores for Pooled Dataset

	Benchmark Model	Addition of Balance of Payments	Addition of Economy Size	Addition of Fiscal Balance	Addition of Reserve Ratio	Addition of Population Growth
Net National Income Per capita Growth	-1.276465 *(1.015851)*	-1.238743 *(1.073819)*	-.1848731 *(1.059207)*	-1.983559 *(1.26891)*	-.9216723 *(1.009284)*	-1.484027 *(1.037796)*
Financial Depth	.2824232 (***) *(.0746427)*	.3020565 (***) *(.0768603)*	.7130838 (***) *(.0881368)*	.177057 (**) *(.0863374)*	.3178618 (***) *(.0752868)*	.2749392 (***) *(.0748664)*
Lagged Inflation rate	-4.529978 (***) *(.7951774)*	-4.947502 (***) *(.8308559)*	-5.533541 (***) *(.8559911)*	-5.830487 (***) *(.9881411)*	-4.360186 (***) *(.7974805)*	-4.549089 (***) *(.7959711)*
Trade Openness	-.4735235 (***) *(.0830356)*	-.4709006 (***) *(.0847333)*	-.9907931 (***) *(.1140931)*	-.4595865 (***) *(.0886242)*	-.608944 (***) *(.0905343)*	-.4686387 (***) *(.0843237)*
Broad Money Growth	-.1454676 *(.2944368)*	-.1512022 *(.298753)*	-.2891343 *(.3095322)*	-.1103645 *(.3291935)*	-.1892656 *(.2988465)*	-.0854684 *(.2948959)*
Balance of Payments		-.3232578 *(.5010389)*				
Economy Size			-18.21294 (***) *(2.435174)*			
Fiscal Balance				2.888767 (***) *(1.007437)*		
Reserve Ratio					.7696859 (***) *(.2082513)*	
Population Growth						-5.34605 (*) *(3.058058)*
Obs	1,848	1,754	1,848	1,404	1,847	1,843
Pseudo R^2	0.1925	0.1978	0.1806	0.1940	0.1890	0.1920

Note: Bootstrapped standard errors (based on 500 replications of the data) are reported in parenthesis. *, ** and *** indicate the significance level of 10%, 5% and 1% respectively

Table 5 – Estimated Propensity Scores for Developing Countries

	Benchmark Model	Addition of Balance of Payments	Addition of Economy Size	Addition of Fiscal Balance	Addition of Reserve Ratio	Addition of Population Growth
Net National Income Per capita Growth	-2.496048 (**) *(1.138479)*	-3.059437 (**) *(1.213007)*	-2.258403 (*) *(1.156711)*	-2.101564 *(1.579864)*	-2.444017 (**) *(1.1463)*	-4.169082 (***) *(1.19454)*
Financial Depth	.9000144 (***) *(.1362396)*	.9052277 (***) *(.1390857)*	.9751324 (***) *(.1497902)*	1.036778 (***) *(.1754861)*	.8991598 (***) *(.1361512)*	.9870625 (***) *(.1399765)*
Lagged Inflation rate	-5.405551 (***) *(.9714361)*	-6.01864 (***) *(1.02634)*	-5.510677 (***) *(.9787865)*	-10.35864 (***) *(1.473679)*	-5.345645 (***) *(.9816258)*	-5.830281 (***) *(1.006583)*
Trade Openness	-.8185527 (***) *(.1224792)*	-.7938969 (***) *(.1243255)*	-.8838519 (***) *(.1340284)*	-1.044121 (***) *(.1511574)*	-.8307571 (***) *(.1277584)*	-.9843568 (***) *(.1349017)*
Broad Money Growth	-.8928326 (**) *(.4051615)*	-.8148823 (**) *(.4076072)*	-.8929745 (**) *(.4052526)*	-.7160013 *(.5049851)*	-.8940103 (**) *(.4056996)*	-.7748029 (*) *(.4097027)*
Balance of Payments		-.9159817 *(.6030573)*				
Economy Size			-4.899757 *(4.023726)*			
Fiscal Balance				-5.323587 (***) *(1.859287)*		
Reserve Ratio					.0943954 *(.2715707)*	
Population Growth						-24.13777 (***) *(4.295561)*
Obs	1,169	1,130	1,169	771	1,168	1,165
Pseudo R^2	0.1955	0.1992	0.1968	0.2055	0.1953	0.1977

Note: Bootstrapped standard errors (based on 500 replications of the data) are reported in parenthesis. *, ** and *** indicate the significance level of 10%, 5% and 1% respectively

Table 6 – Estimated Propensity Scores for Industrial Countries

	Benchmark Model	Addition of Balance of Payments	Addition of Economy Size	Addition of Fiscal Balance	Addition of Reserve Ratio	Addition of Population Growth
Net National Income Per capita Growth	1.13468 *(2.502828)*	3.487402 *(2.670006)*	7.270026 (**) *(2.92465)*	2.231877 *(2.685675)*	.5784913 *(2.503518)*	.9804571 *(2.589494)*
Financial Depth	.4378259 (***) *(.1230879)*	.507338 (***) *(.1314425)*	1.169334 (***) *(.1604832)*	.5088195 (***) *(.1297249)*	.357966 (***) *(.1264959)*	.3106425 (**) *(.1287727)*
Lagged Inflation rate	-5.296096 (***) *(1.940595)*	-4.515596 (**) *(1.967882)*	-9.29835 (***) *(2.223743)*	-4.333513 (**) *(2.050287)*	-5.866962 (***) *(1.952283)*	-6.76392 (***) *(2.013171)*
Trade Openness	-.3956055 (***) *(.1211312)*	-.5184551 (***) *(.1385288)*	-2.060005 (***) *(.314728)*	-.6238583 (***) *(.1365999)*	-.7416269 (***) *(.169846)*	-.700291 (***) *(.1324769)*
Broad Money Growth	.5185425 *(.5334782)*	.5816907 *(.543496)*	.1241169 *(.5911033)*	.004443 *(.5576803)*	.5632188 *(.5276614)*	.0205321 *(.5650736)*
Balance of Payments		1.654857 *(1.167891)*				
Economy Size			-27.8596 (***) *(3.697125)*			
Fiscal Balance				8.395763 (***) *(1.559049)*		
Reserve Ratio					1.711032 (***) *(.6017659)*	
Population Growth						57.53441 (***) *(10.56915)*
Obs	679	637	679	633	679	678
Pseudo R^2	0.1832	0.1897	0.2263	0.1821	0.1752	0.1881

Note: Bootstrapped standard errors (based on 500 replications of the data) are reported in parenthesis. *, ** and *** indicate the significance level of 10%, 5% and 1% respectively

6 Matching Methods

The propensity score has been a frequently used tool in research where bias removal is a prerequisite and as such is widely used across various research fields (Chih-Lin, 2011). The method itself undergone a rapid development in the last decade with many researchers proposing a variety of propensity score based evaluation methods. In this section I provide a general overview of all matching methods introduced in this research.

6.1 Nearest Neighbor Matching

Nearest Neighbor is the most commonly implemented matching algorithm that conducts one at a time matching of treatment and control subjects. Under this method the treated subjects are first randomly sorted and then the first treated subject is paired with n control subjects such that the difference in their propensity scores is minimized[1], i.e. the *Nearest Neighbor*. After the first matched subjects are removed from the pool, the matching continues for the rest (Gu & Rosenbaum, 1993). Once each of the treated subjects has been paired with n control subjects, the difference between the outcomes of the matched subjects is calculated. The ATT is then computed by averaging the calculated differences (see Equation 5.7).

For our purposes, the method is applied both with (for $n = 1$, i.e. *Nearest Neighbor*) and without the no-replacement condition[2] (for $n = 3$, i.e. *3 Nearest Neighbor*). When using matching without replacement, once a control subject has been matched to a treated subject it is removed from the pool and is no longer available for subsequent matching (Austin, 2011b). Allowing for replacement can be useful with data where distribution of the propensity score among treatment and control group varies significantly. With such data, high-score subjects might be matched to low-score control subjects and the quality of matching may therefore struggle. Allowing for replacement should lead to reduction of number of distinct control subjects used to construct the counterfactual outcome and thereby increase quality of the matching (Smith & Todd, 2005).

---

[1] If several control subjects have equal propensity score, one of the control subjects is selected and paired with the treated subject on random (Austin, 2011b)

[2] No-replacement condition requires that a control subject can be a best match for only one treated unit.

6.2 Radius Matching

As discussed above, allowing for replacement might solve some risk of bad matching for the *Nearest Neighbor* matching method. Notwithstanding that, the quality of matching might still be negatively influenced if the closest control neighbor is far away from the treated subject. One way to incorporate such risk is to impose a tolerance on the maximum distance in propensity scores between treated and control subjects, i.e. *caliper*. This in some sense works similar to replacement condition for Nearest Neighbor and should provide comparable elimination of bad matches. Setting a caliper means that treated subject can be paired only with control subjects that are close enough to its propensity score such that the caliper condition holds. If there is no close enough neighbour for the treated subject to be paired with, it is eliminated from the sample and not included in calculation of the ATT.
Shortcoming of this method is that there is no *a priori* correct level for a caliper setting[3] (Smith & Todd, 2005). If we decide to set calliper too tight, majority of treated subjects might get eliminated from the matching just because there are no control subjects to be paired with that would satisfy the caliper setting. On the contrary if we set caliper too loose the method might not be able to filter out bias caused by far away control subjects (bad matches). In order to incorporate this fact, I introduce four different caliper setting. Specifically, a wide caliper $r = 0.03$, a medium caliper $r = 0.01$, a tight caliper $r = 0.005$ and an extra tight caliper $r = 0.001$. If we assume for example a caliper of $r = 0.005$, this setting will allow the treated subject to be paired only with a control subject that is in its proximity by P_i which is lower than, or equal to r; $r = 0.005$, hence the following has to hold:

$$P_i \leq r\ ;\ P_i = [abs(PS_t) - abs(PS_c)] \tag{6.1}$$

Where, PS_t is the propensity score of treated subject and PS_c is the propensity score of control subject.

---

[3] Cochran and Rubin (1973) assessed reduction in bias using various calipers with widths proportional to the standard deviation of the confounding variable. They suggest to use caliper proportionate to the standard deviation (Cochran & Rubin, 1973). Optimal caliper widths for estimating risk differences and differences in means were assessed by Austin (2011a). Austin suggests use of a caliper of width equal to 0.2 of the standard deviation of the propensity score. This value should minimize the mean squared error of the estimated treatment effect (Austin, 2011a).

For purpose of this thesis I use a *radius* version of caliper matching. This method matches not only the nearest neighbor within the caliper but also all other comparable control subjects that lie within the specified area (Dehejia & Wahba, 2002). The method therefore allows for use of some extra control subjects when good enough matches are available[4].

6.3 Kernel Matching

Both Nearest Neighbor and Radius matching have in common that only a few control observations are used to construct the counterfactual outcome for treated subjects. Kernel matching on the other hand represents a non-parametric estimator that uses weighted averages of all subjects in the control group to construct the counterfactual outcome.

Since Kernel matching uses all untreated observations to compute the counterfactual it is important to incorporate region of common support to apply this method correctly. This ensures that any combination of characteristics observed in the treatment group can also be observed among the control group (Bryson, et al., 2002). In order to comply, all control observations with propensity score higher than the highest or lower than the lowest propensity score of treated subjects are eliminated from the analysis and are not included in the calculation of ATT[5]. This approach should eliminate possible bias caused by bad matches (Heckman, et al., 1997).

Since the counterfactual are estimated based on the whole control sample this approach provides lower variation of estimates. On the other hand, if the imposition of area of common support is improper, estimates might be biased due to bad matches (Heckman, et al., 1997).

Smith & Todd (2005) note that kernel matching can be seen as a weighted regression of the counterfactual outcome on an intercept with weights given by the kernel weights. The weight depends on the distance between each control subject and the treatment subject. The average puts higher weights on subjects that are close to the treated subject in terms of the propensity score and lower weight on subjects that are more distant (Caliendo & Kopeinig, 2005). The asymptotic distribution of Kernel estimator is in detail derived in Heckman, et al. (1998).

[4] Hence combines features of replacement setting for Nearest Neighbor method while avoiding risk of bad matching (Caliendo & Kopeinig, 2005).

[5] Also known as *minima maxima*, further discussed in section 7.2.1. M*inima maxima* approach is applied for each matching method and for each dataset.

When applying kernel matching one has to choose the kernel function and the bandwidth parameter. The first point appears to be relatively unimportant in practice (Dinardo & Tobias, 2001). Bandwidth on the other hand is seen as much more important attribute (Caliendo & Kopeinig, 2005) because it provides us with the following trade-off: High bandwidth-values yield a smoother estimated density function, leading to a better fit and decreased variance between estimated and true density function. On the other hand, it might smooth-away underlying features and therefore cause biased estimates. The bandwidth choice is therefore a compromise between a small variance and an unbiased estimate (Caliendo & Kopeinig, 2005).

Following Lin & Ye (2007, 2009, 2012) and Lin (2010) I conduct ATT estimation using Epanechnikov[6] kernel function with a default bandwidth of 0.06.

7 Model

In this section I use propensity score to estimate the overall ATT effect of IT adoption in pooled, industrial, and developing datasets by implementing all matching techniques presented in section 6. The outcome variables of our interest are inflation variability, inflation, real per capita GDP variability, and real per capita GDP growth. Quality assessment of the matching is presented in section 7.2.

7.1 Outcome Estimation

Here, I conduct the ATT estimation based on propensity score from section 5.4, using all matching methods that were presented in section 6. ATT is estimated in STATA software, using PSMATCH2 module developed by Leuven & Sianesi (2003)[7]. ATT estimates for the benchmark model are presented in the Table 7.

---

[6] $K(u) = \frac{3}{4}(1 - u^2)1_{\{|u| \leq 1\}}$

[7] PSMATCH2 module implements full Mahalanobis matching and a variety of propensity score matching methods to adjust for treatment effect assessment (Leuven & Sianesi, 2003).

Table 7 - ATT estimates for Benchmark model

Outcome Variable	Dataset	Matching Method							
		Nearest Neighbor	3 Nearest Neighbor	Radius Matching r = 0.03	Radius Matching r = 0.01	Radius Matching r = 0.005	Radius Matching r = 0.001	Kernel Matching	IPW
Inflation Variability	Pooled	-.001813897 (**)	-.002196013 (**)	-.003235252	-.003134929	-.002846847 (**)	-.002314619 (**)	-.00225596 (**)	-.00237162 (***)
		(.000883088)	*(.001019554)*	*(.003872085)*	*(.003735423)*	*(.001336256)*	*(.001090309)*	*(.001116505)*	*(.00076421)*
	Developing	-.005666963 (***)	-.008285733 (***)	-.007778182	-.00796234	-.007692291 (***)	-.010779938 (***)	-.009547059 (***)	-.030584 (***)
		(.001759446)	*(.003087618)*	*(.006800618)*	*(0.006722966)*	*(.002254033)*	*(.003383079)*	*(.00254033)*	*(.0106138)*
	Industrial	-.001567135	-.000494315	-.000115123	-.000376806	-.000574636	-.000670167	-.000017553	-.00048784
		(.001958093)	*(.001054701)*	*(.001102266)*	*(.001175389)*	*(.001252063)*	*(.001168323)*	*(.001609769)*	*(.00037628)*
Inflation	Pooled	-.000255963	-.000181001	.000260521	.000340882	.00036946	.000354828	-.000785106	-.00011812
		(.000993626)	*(.001031591)*	*(.000975873)*	*(.000970405)*	*(.000956659)*	*(.00099632)*	*(.001271435)*	*(.00008066)*
	Developing	-.000201121	-.001013666	-.000497142	-.000819686	-.001662806	-.003710111 (**)	-.001830086	-.0012879
		(.001473233)	*(.001500434)*	*(.00146247)*	*(.001467745)*	*(.001429171)*	*(.001654854)*	*(.001990413)*	*(.0013356)*
	Industrial	-.000071124	-.000927556	-.000199785	-.000383409	-.000520402	-.001539417 (*)	-.000774677	-.008417
		(.001415583)	*(.000927146)*	*(.000996062)*	*(.001043555)*	*(.00110089)*	*(.00097397)*	*(.001136078)*	*(.0011407)*
Real per capita GDP Variability	Pooled	-.001513437	-.001754038	-.001432745	-.001626789	-.001697628	-.002412899 (*)	-.001767355	-.00014775
		(.001238048)	*(.001467499)*	*(.001142057)*	*(.001147591)*	*(.001167404)*	*(.001310737)*	*(.001526329)*	*(.009519)*
	Developing	-.002626555	-.003875753 (*)	-.002592488	-.003346116	-.002558918	-.00504455 (**)	-.003476099	-.0028767
		(.002449874)	*(.002225698)*	*(.001767104)*	*(.001815015)*	*(.00183803)*	*(.00246806)*	*(.003488426)*	*(.0018111)*
	Industrial	-.001250008	-.001434575	-.001497256	-.001484321	-.001322317	-.001462183	-.001226911	-.0012802
		(.001309296)	*(.001355545)*	*(.001206739)*	*(.001248301)*	*(.001280055)*	*(.001449767)*	*(.001599404)*	*(.0011195)*
Real per capita GDP Growth	Pooled	.000370776	.000673394	.000883278	.001247065	.001699038	.002337291	.000391929	.0014176
		(.002396511)	*(.002364077)*	*(.002024123)*	*(.002050875)*	*(.002066231)*	*(.002225962)*	*(.002940168)*	*(.0017575)*
	Developing	.001845081	.001312623	.000976702	.000360411	.000014951	.003616766	.003616766	.0013042
		(.003603256)	*(.00346427)*	*(.002998158)*	*(.003044988)*	*(.003063738)*	*(.003775361)*	*(.003775361)*	*(.0037214)*
	Industrial	.003048638	.001537141	.002413005	.003077445	.0021591	.001674252	.003882323	.004022 (*)
		(.002595765)	*(.002746947)*	*(.002412613)*	*(.00252868)*	*(.002618637)*	*(.002967115)*	*(.003226636)*	*(.0022922)*

Note: IPW stands for Inverse-Probability Matching – see section 8.2 for more detail, bootstrapped standard errors (based on 500 replications of the data) are reported in parenthesis, *, ** and *** indicate the significance level of 10%, 5% and 1% respectively

Concerning inflation variability, for all three datasets the ATT estimates are in all cases negative. It appears that on average, IT countries exhibit lower inflation variability than countries with other monetary regimes. For pooled and developing dataset the estimates are also in most cases statistically significant (apart from Radius matching with looser caliper settings), however they remain insignificant for industrial countries. This might be caused by proportionately smaller industrial data sample. The estimates for pooled and developing datasets are similar among all matching methods, but they are more variable for industrial dataset. This might be again due to smaller data sample. In overall, tighter calipers (*r=0.005; r=0.001*) provide us with statistically more significant results compared to loose calipers *(r=0.03, r=0.01).* This might be due to the fact that tighter caliper filters out far away control subject that would otherwise represent bad matches. The results are consistent with Lin & Ye (2007) and Ardakani, et al. (2015).

In terms of inflation, the estimates are all negative for developing and industrial datasets. It seems that on average IT countries are subject to lower inflation rates than countries with other monetary regimes. Estimates for pooled dataset are on the other hand somewhat mixed. While radius estimates are positive, the rest is negative. This is probably caused by some bad matching between industrial and developing countries and it only emphasizes how important it is to clearly distinguish between industrial and developing countries. The only two statistically significant estimates in developing and industrial datasets are on account of extra tight caliper *(r=0.001)*. Overall, estimates are not as stable and exhibit higher variation than estimates for inflation variability. The results are consistent with Lin & Ye (2007, 2012) and Ardakani, et al. (2015),

Situation is similar for real per capita GDP variability. All estimates for each dataset have negative properties. Hence it appears that IT countries exhibit lower GDP variability than non-targeters. Estimates are however only scarcely statistically significant. Three exceptions are the estimates for pooled and developing dataset under extra tight caliper and the estimate for developing dataset under *3 Nearest Neighbor*. Notwithstanding that, the results are much more stable and exhibit lower variability than coefficients estimated for inflation rates.

Statistically insignificant are also results for GDP growth. All estimates have the same positive properties which indicates that IT countries experience higher growth rates of GDP compared to non-targeters. On the other hand, the results are much less stable and more variable than estimates for GDP variability. The results are partly consistent with Lin & Ye (2012) [82]

To summarize, the results indicate that both industrial and developing inflation targeting countries experience lower inflation levels and at the same time higher output growth than non-targeting countries. Moreover, it appears that both industrial and developing countries achieve combination of lower inflation variability and outcome variability compared to non-targeting countries. Nonetheless it is important to point out that majority of the estimates are statistically insignificant.

7.2 Matching Quality Assessment

In this section, I assess quality of the matching performed in the previous section. In an observational study the true propensity score is not known and must be estimated using the available dataset. It is therefore important to assess whether or not the propensity score model has been adequately specified and whether the propensity score is a *balancing score*, i.e. whether the distribution of our benchmark covariates is independent of treatment assignment (Heckman et al, 1997; Austin, 2001b).

Naturally, the most appealing way to assess the balancing properties would be to use a t-test to check for statistical significance of difference in covariate means for both treated and control groups (Rosenbaum & Rubin, 1985). This approach; however, has been broadly criticized and discouraged for its tendency to produce misleading results. Since significance levels are confounded with the sample size and the matched sample is often inevitably smaller than the original one (due to elimination of not matched observations), the significance testing is likely to be biased (Austin, 2008)[83]. Moreover, the t-test is not able to provide us with information on the bias reduction before and after the matching and therefore tells us nothing about quality of the matching (Caliendo & Kopeinig, 2005).

---

[82] Lin & Ye (2012) exhibits mixed properties of estimates for real per capita GDP growth. My results are consistent under caliper and kernel methods.

[83] Furthermore, Imai, et al. (2008) suggested that balance is a property of a particular sample and that reference to *super population* is inappropriate (Imai, et al., 2008).

7.2.1 Standardised Bias

Assessing standardised differences is a widely recommended alternative to t-test when examining balancing properties of the propensity score and the PSM model (Caliendo & Kopeinig, 2005). Since we do not condition on all covariates but only on the propensity score, it is crucial that the propensity score balances the distribution of the relevant covariates in both treated and control groups – i.e. *conditional independence assumption* (see part 5.2 for detail). In order to ensure that the propensity score has been estimated properly, the distribution of relevant covariates among matched sample has to be similar in both treated and control groups. Significant differences in the covariates' distribution after the matching might signal violation of the *conditional independence assumption* and lead to biased estimates (Austin, 2011b).

To determine the level of similarity of treated and control group in the matched sample it is recommended to assess the means or medians of continuous covariates and their distribution between treated and control subjects (Flury & Riedwyl, 1986; Austin, 2009b). For a continuous covariate, the standardized difference is defined as (Austin, 2009b, p. 2039):

$$d = \frac{(\bar{x}_t - \bar{x}_c)}{\sqrt{\frac{V_t^2 + V_c^2}{2}}} \tag{7.1}$$

where $\bar{x}_t$ is the sample mean of the covariate for treated subjects, $\bar{x}_c$ is the sample mean of covariates for control subjects, S_t^2 is the sample variance of the covariate for treated subjects and S_c^2 is the sample variance of the covariate for control subjects.

In randomized controlled experiment, randomization ensures that, on average, treated subjects do not systematically differ from control subjects in both measured and unmeasured characteristics. Under such perfect conditions the standardised difference in covariates' distribution would equal or be close to zero. Aim of the PSM is to mimic such conditions[84]. Since we condition only on the propensity score, we have to check if the matching procedure is able to balance the distribution of the relevant covariates in both the control and treatment group, and therefore whether our propensity score is a *balancing score* (Caliendo & Kopeinig, 2005).

[84] Considering equation 5.5, we are concerned only with the observed characteristics.

Most commonly used indicator to assess the marginal distribution of the X-variables is the standardised bias (hereinafter "SB") proposed by Rosenbaum & Rubin (1985) [85]. The basic idea of the approach is to compare the situation before and after matching and check if there remain any major differences after conditioning on the propensity score. If the major differences remain, the matching was not completely successful and the estimates might be biased (Caliendo & Kopeinig, 2005).

For each covariate X the SB is defined as the difference of sample means in the treated and matched control subsamples as a percentage of the square root of the average of sample variances in both groups (Caliendo & Kopeinig, 2005). Taking into account equation 7.1, the SB before matching is given by (Caliendo & Kopeinig, 2005, p. 15):

$$SB_{before} = 100(\frac{(\bar{X}_t - \bar{X}_c)}{\sqrt{0.5(V_t(X) + V_c(X))}} \tag{7.2}$$

And after matching as (Caliendo & Kopeinig, 2005, p. 15):

$$SB_{after} = 100(\frac{(\bar{X}_{tM} - \bar{X}_{cM})}{\sqrt{0.5(V_{tM}(X) + V_{cM}(X))}} \tag{7.3}$$

where X_t and V_t is the mean and variance in the treatment group before matching and X_c and V_c the analogue for the control group. X_{tM}, V_{tM}, X_{cM} and V_{cM} are the corresponding values for the matched samples (Caliendo & Kopeinig, 2005). A possible problem with the SB is that there is no universally accepted consensus on threshold for identification of successful matching. Some empirical studies consider bias reduction below 10 per cent as "sufficient" (Normand, et al., 2001; Imai, et al., 2008). Other studies however recommend reduction under 5 per cent (Caliendo & Kopeinig, 2005). In my assessment I reflect both of the proposed thresholds.

SB assessment for all three datasets and all matching techniques is presented in tables 22 - 28 in the appendix. Graphical analysis of the same is presented by figures 4 - 24 in the appendix. The matching appears to be the most successful for developing countries for which

[85] For example, Lechner (1999), Sianesi (2004), and Caliendo, et al. (2005).

the mean bias peaks at 4.9 per cent and never exceed the 5 percentage threshold. Compared to unmatched sample it reduces the bias between control and treatment group usually by around 90 per cent which is a great enhancement. However, on individual basis, in five cases the financial depth bias slightly exceeds the 5 percentage threshold, in one case even the 10 percentage threshold and peaks at 13.1 per cent for *Nearest Neighbor* matching. In this case the propensity score might not meet the *conditional independence assumption* and the ATT estimate may therefore be biased.

On the mean basis, the matching also appears to be successful for industrial dataset for which the mean bias does not exceed 5.8 per cent. Again, compared to the unmatched sample, matching in all cases succeeds in reducing the bias between control and treatment group. However, individual values of financial depth appear to be systematically biased by more than 10 per cent for all matching methods. This may adversely affect quality of the ATT estimates.

Mean bias appears to be also very stable for pooled dataset, which exceeds the 5 percentage threshold only once and peaks at 5.3 per cent. For rest of the cases the mean bias remains under the 5 percentage threshold. Compared to industrial and developing dataset, there are no individual extremes that would exceed 10 percentage bias. This might be thanks to larger data sample. On the other hand, the individual biases tend to exceed the 5 percentage threshold more often than in case of developing and industrial dataset. This might very well go again on the account of large data sample in which we match industrial and developing countries with dissimilar characteristics.

7.2.2 Area of Common Support

Ho, et al. (2007) suggest that the SB assessment should be complemented by comparing the distribution of the estimated propensity score between treated and control subjects in the matched sample. This approach might be useful for determining the common support area and the degree of overlap in the propensity score between treated and control subjects. Furthermore, it may provide a rough assessment of whether the means of covariates are similar between the two groups (Austin, 2011b).

In section 5.2 I discuss that *area of common support* has to be properly established to ensure proper propensity score estimation and unbiased matching. Area of common support ensures that both treated and control subjects in the matched sample have to some extent similar characteristics in terms of their propensity score. It might be referred to as setting boundaries for minimum and maximum propensity score values for control subjects in the matched sample. Control observations with propensity score below the minimum or above the maximum are eliminated from the matched sample and are not used for computation of the ATT estimates.

Dedication of *common support area* is especially important when applying Kernel Matching, under which all untreated observations are used to compute the missing counterfactual[86]. Not establishing the area of common support could cause the model to perform some bad matches on control subjects below or above the boundary. This might result to biased ATT estimates (Caliendo & Kopeinig, 2005; Austin, 2011b).

I establish area of the common support, using *minima maxima* comparison for the benchmark model and each of the datasets. All control observations with propensity score higher than the highest or lower than the lowest propensity score of treated subjects are eliminated from the analysis and are not included in the calculation of ATT. This should eliminate a source of possible ATT bias (Caliendo & Kopeinig, 2005). Common support area for each dataset is numerically expressed in Table 8.

Table 8 - Area of Common Support, benchmark model (*propensity score*)

	Pooled dataset	Developing dataset	Industrial dataset
Max propensity score	0,4540	0,7397	0,5433
Min propensity score	0,0684	0,1005	0,1791

Figure 1, Figure 2, and Figure 3 depicts Benchmark model's propensity score distribution for pooled, developing, and industrial dataset respectively. The figures allow us to illustrate the area of common support graphically. Control observations that did not comply with *minima maxima* and were eliminated from the matching pool are depicted in green.

It is worth noting that a weak spot of the *minima maxima* approach lies around the boundaries of the common area. It is probable that while eliminating control observations with propensity scores lower than the lowest or higher than the highest treated subject, we eliminated also several control observations that were very close to the boundary and that would otherwise represent a good match. This might have some adverse effect on the ATT estimates.

[86] Same applies for *Inverse Probability Weighting* (see section 8.2 for detail on the *IPW*)

7.2.3 Overlap

In the previous sections I assessed similarity of treated and control group in the matched samples based on assessment of *standardised bias,* furthermore I conducted *maxima minima comparison* to establish area of common support. In this part, I aim to extend the analysis by another supplemental testing.

Literature concerning PSM suggests that the similarity of benchmark covariates in the matched sample for control and treated group may be complementary assessed graphically, by comparing distribution of their estimated propensity score. (Austin, 2011b; Ho, et al., 2007). While this approach, standalone, would not be sufficient to determine whether the propensity score estimation is not biased, it is useful for determining the degree of overlap between treated and control groups (Austin, 2009a). With greater overlap, matching tends to provide us with more precise ATT estimates. If there is little or no overlap in the control group, the estimates might be biased due to bad matches (Rubin & Thomas, 1992). Observing the Figure 1, Figure 2, and Figure 3, it is clear that distribution of propensity score among treated and control group is not perfectly similar. In all three cases the overlap is very generous for observations with low and medium propensity scores. However, with increasing scores the overlap starts to decay. This fact might be of concern for matching conducted under no-replacement condition (in our case *Nearest Neighbor*). Under these circumstances it might be likely that treated observations will be matched with control observations whose propensity score are not perfectly similar, because another possibly much better match may have already been eliminated from the pool. In the end, this might provide us with biased ATT estimates. Notwithstanding that, this should not present any problem for the rest of the applied matching methods all of which allow for replacement. Under this setting, the control subjects are not eliminated from the pool once they are matched, on the contrary they can be used for multiple matching. Moreover, the overlap appears to be in all three cases distributed widely-enough and across the whole sample population. We can therefore expect that estimates for matching methods that incorporate replacement option should not suffer from any estimate bias that would go on the account of insufficient overlap or undedicated area of common support.

Figure 1 - Propensity score distribution (density), Benchmark model, Pooled dataset

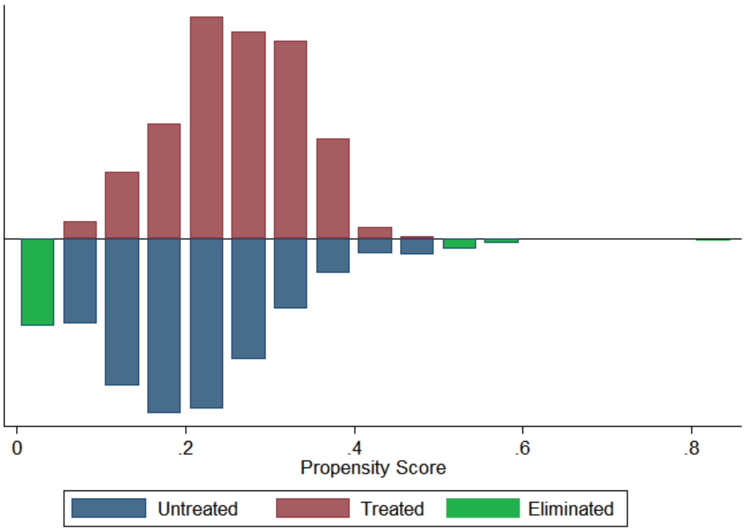


Figure 2 - Propensity score distribution (density), Benchmark model, Developing countries

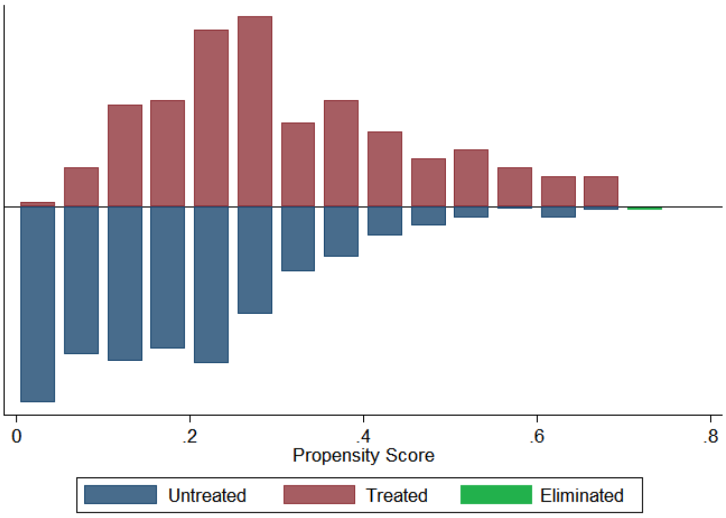
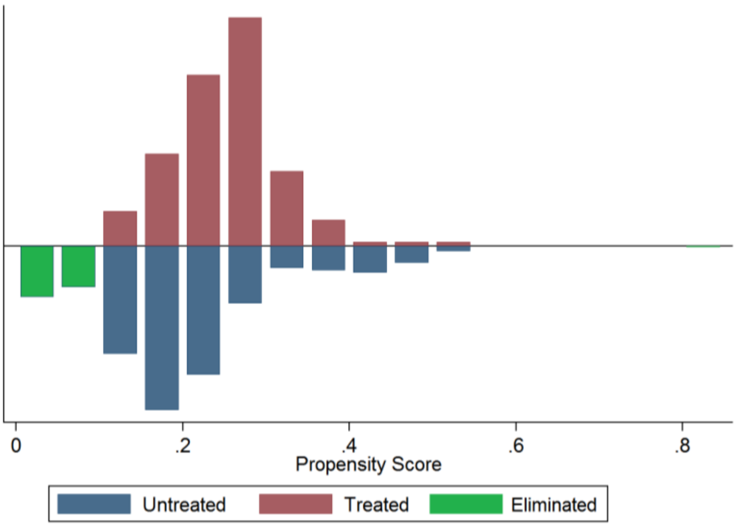


Figure 3 - Propensity score distribution (density), Benchmark model, Industrial countries



8 Robustness Testing

In section 4.2 and 4.3 I discuss methodology applied for a country's assignment into either developing or industrial dataset, and its identification as an IT or non-IT country. In section 7.1, I present ATT estimates that are based on this benchmark assignment. Here, I conduct robustness testing of those properties, first by altering the countries' aforementioned assignments and conducting the ATT estimation all over again. Afterwards I introduce the IPW as a substitute to PSM methods.

8.1 Alternative IT Adoption & Group Assignment

For the benchmark model, the countries are assigned to either developing or industrial dataset based on J.P. Morgan's EMBI, FTSE's Annual Country Classification Review and IMF's World Economy Outlook. There are several discrepancies between the sources. While FTSE classifies the Czech Republic and Israel as emerging countries, EMBI and IMF accounts both of them as Industrial. Similar applies for the Korean Republic and Slovenia, which are considered as industrial by IMF. It is the other way around for Greece, which is considered industrial by IMF but as emerging by EMBI. To check for possible misclassification, I conduct robustness testing by altering the aforementioned assignment. For this purpose, I reclassify the Czech Republic, Israel, Korean Republic, and Slovenia as industrial countries[87] and Greece as a developing country[88]. See column Robustness check in Table 11 (appendix) for details.

A country's identification as inflation targeter and its IT adoption dates for the benchmark model are obtained from Rose (2007), Miao (2009), and Hammond (2012). There are however certain discrepancies in their methodologies. This leads to a situation in which one source accounts a country as an IT while the other two sources don't. To check for possible misclassification, I conduct robustness testing by altering some of the controversial assignments. Specifically, I reclassify Armenia, Indonesia, Romania, Serbia, and Switzerland as non-targeters. This shortens our list of inflation targeters by 4 for developing

---

[87] Czech Republic for periods 2008-2015, South Korea for periods 2007-2015, Slovenia for periods 2003-2015 and Israel for periods 1990-2015. These are the breaking points, when real per capita GDP of each country exceeded 20,000 USD.

[88] For all observations

dataset and by 1 for industrial dataset. See column Robustness check in Table 12 (appendix) for details.

8.1.1 Adjusted Propensity Score Estimation

In this section I conduct the first stage estimation of propensity score for all three datasets after the robustness testing adjustments. I use a probit regression introduced in section 5.3. In all cases the dependant variable is the dummy variable for inflation targeting. Outcome-independent variables used for the estimation are discussed in detail in section 5.4. The estimates are presented in Table 9.

Table 9 - Adjusted propensity score estimates for Benchmark model

Benchmark model adjusted for robustness testing			
	Pooled dataset	**Developing countries**	**Industrial countries**
Net National Income Per capita Growth	-1.634252 *(1.066384)*	-2.777859 (**) *(1.190548)*	.9109404 *(2.527977)*
Financial Depth	.3242868 (***) *(.0767826)*	1.077095 (***) *(.1386259)*	.052612 *(.1211199)*
Lagged Inflation rate	-4.872878 (***) *(.8600733)*	-6.115206 (***) *(1.086705)*	-4.563418 (***) *(1.946936)*
Trade Openness	-.5253664 (***) *(.0867478)*	-.7330197 (***) *(.1256447)*	-.4997896 (***) *(.1242514)*
Broad Money Growth	-.1313497 *(.3020389)*	-.7236759 (*) *(.4169824)*	.4475503 *(.5051513)*
Obs	1,848	1,136	712
Pseudo R2	0.1657	0.1731	0.1615

*Note: Bootstrapped standard errors (based on 500 replications of the data) are reported in parenthesis. *, ** and *** indicate the significance level of 10%, 5% and 1% respectively*

Properties of the estimates remained in all cases the same as for the benchmark model. Furthermore, all estimates are highly similar to the benchmark estimates[89] (see section 5.4). It appears that the reclassifications did not have significant effect on the estimates or their quality.

8.1.2 Adjusted ATT Estimation

In this section, I conduct ATT estimation using adjusted propensity score estimates from the previous section and all matching methods that were presented in section 6. ATT is again

[89] Apart from financial depth for industrial countries, which is no longer statistically significant.

estimated in STATA software, using PSMATCH2 module developed by Leuven & Sianesi (2003). The estimates are presented in Table 10.

Compared to the benchmark model, there are several differences in the estimated outcomes. In terms of Inflation variability, the model performs equally well or better than the benchmark. The estimates have the same properties and in some cases they are even more statistically significant than before (*Nearest Neighbor, 3 Nearest Neighbor,* and *Kernel* under industrial dataset). Standard errors appear to be similar to benchmark as well.

Concerning Inflation, compared to the benchmark, the adjusted model appears to clarify estimates for pooled dataset which are now negative in all cases, yet still statistically insignificant. On the other hand, *radius* (0.001) estimates for developing and industrial dataset are not statistically significant anymore.

The similar applies for GDP and GDP variability. Estimates have in all cases the same properties as the benchmark ones but majority of them remains statistically insignificant.

To sum it up, it appears that conducted adjustments did not have any major impact on the estimates. In some cases, the adjusted model performs better than the previous one but in other cases it is the other way around. Nevertheless, it is worth noting that the adjusted model managed to clarify contradictory estimates for IT adoption effect on Inflation under pooled dataset. All of the mentioned estimates now have negative properties.

Standardise bias assessment for all three datasets and all matching techniques is presented in tables 29 – 35 in the appendix. The matching appears to be again the most successful for developing countries for which the mean bias peaks at 4.8 per cent and never exceed the 5 percentage threshold. Individual bias peaks at 9.5 per cent for financial depth. Compared to unmatched sample it reduces the bias between control and treatment group usually by around 80-95 per cent which is a great enhancement that is comparable to the benchmark model.

On the mean basis, the matching is much less successful for industrial dataset. The mean bias does not exceed 10 percentage threshold however it does not succeed in falling under 5 per cent either. The individual peak is at 11.2 per cent for financial depth. Furthermore, under *Nearest neighbor* the matching does not succeed in reducing bias for net national income. This does not have to be interpreted as violation of conditional independence assumption however the goal of matching is to reduce this bias which clearly did not happen. This may adversely affect quality of the specific ATT estimate.

Table 10 - ATT estimates, Robustness testing adjusted model

Outcome Variable	**Dataset**	**Matching method**							
		Matching Methods							
Variable	Dataset	Nearest Neighbor Matching	3 Nearest Neighbour Matching	Radius Matching				Kernel Matching	IPW
				r = 0.03	r = 0.01	r = 0.005	r = 0.001		
Inflation Variability	Pooled	-.00506719 (***) *(.00174448)*	-.003722102 (**) *(.001520348)*	-.003493696 *(.003838185)*	-.003401736 *(.003724834)*	-.003304711 (**) *(.001409138)*	-.003530052 (*) *(.001845361)*	-.005881241 (**) *(.002412797)*	-.0046579 (***) *(.001576497)*
	Developing	-.005393289 (***) *(.001362528)*	-.006829443 (***) *(.002155646)*	-.006720531 *(.006539869)*	-.006246384 (***) *(.002274836)*	-.006445184 (**) *(.002539032)*	-.008386276 (***) *(.00230026)*	-.005690009 (***) *(.001721005)*	-.019723(**) *(.0104795)*
	Industrial	-.00189149 (***) *(.000868703)*	-.000809935 *(.001105111)*	-.000909267 *(.00088237)*	-.00095903 *(.000934293)*	-.001063203 *(.000939091)*	-.000944971 *(.001015346)*	-.001725819 (*) *(.00100434)*	-.00051247 *(.000364897)*
Inflation	Pooled	-.001016125 *(.001208022)*	-.000207166 *(.001137069)*	-.000387942 *(.001008337)*	-.000595899 *(.000972912)*	-.000430393 *(.000975448)*	-.000223072 *(.001107164)*	-.000182805 *(.001548806)*	-.000176549 *(.00013973)*
	Developing	-.000090969 *(.00151784)*	-.000045066 *(.001573989)*	-.000286595 *(.00154629)*	-.000040868 *(.00147897)*	-.000147352 *(.001572544)*	-.002415696 *(.001754855)*	-.001001965 *(.001916475)*	-.0012127 *(.001029798)*
	Industrial	-.001341711 *(.000943603)*	-.001662926 *(.001063607)*	-.001267201 *(.000852744)*	-.001258311 *(.000903217)*	-.001206048 *(.00094029)*	-.00145296 *(.000956096)*	-.001078629 *(.001163602)*	-.009732 *(.008133582)*
Real per capita GDP Variability	Pooled	-.001308136 *(.001133689)*	-.001183877 *(.001097573)*	-.001755507 *(.001103673)*	-.0018691 (*) *(.001112375)*	-.001799987 *(.001133027)*	-.000911579 *(.001296094)*	-.001587775 *(.001327615)*	-.0005497 *(.00385762)*
	Developing	-.002124219 *(.001862637)*	-.003172236 *(.00236036)*	-.00352658 (**) *(.001725536)*	-.003790837 (**) *(.00176483)*	-.00483952 (**) *(.00191136)*	-.001698288 *(.002181352)*	-.002921992 *(.002439862)*	-.0009147 *(.001761435)*
	Industrial	-.000486703 *(.001292772)*	-.000153769 *(.00135763)*	-.000084403 *(.001181076)*	-.000177633 *(.001220015)*	-.000237224 (.001250606)	-.000640184 *(.001331341)*	-.00057093 *(.001460418)*	-.0036549 *(.004254438)*
Real per capita GDP Growth	Pooled	.002802952 *(.002470803)*	.003956528 *(.00226789)*	.002252662 *(.001980888)*	.002771987 *(.001987695)*	.002354379 *(.002017316)*	.002883551 *(.002262275)*	.00466854 *(.002983327)*	.0018604 *(.00201460)*
	Developing	.001984088 *(.003811692)*	.001973239 *(.003552989)*	.000542574 *(.003010633)*	.001870007 *(.003037875)*	.003310561 *(.003252799)*	.00204808 *(.003493789)*	.00427943 *(.004726851)*	.0016471 *(.002853422)*
	Industrial	.002039209 *(.002487693)*	.000756673 *(.002606707)*	.002102327 *(.002216779)*	.001874384 *(.002322855)*	.001698279 *(.002401452)*	.001163778 *(.002583224)*	.002467953 *(.002824802)*	.0038504 *(.00326473)*

Note: bootstrapped standard errors (based on 500 replications of the data) are reported in parenthesis, *, ** and *** indicate the significance level of 10%, 5% and 1% respectively

Pooled dataset is again somewhere half-way between the developing and industrial datasets. Similarly to industrial dataset, the mean bias does not exceed 10 percentage threshold but on the other hand struggles to fall under 5 per cent. Individual bias peaks at 13.5 per cent for net national income. The estimates might therefore be adversely affected.

Based on the afore-performed SB assessment, we can say that the model did not benefit from the robustness adjustments. On one hand there are some improvements in terms of ATT estimates, on the other hand these are not supported by improvements to balancing properties. On the contrary the balancing properties appear to be less stable than in case of the benchmark model. Hence the PSM is less able to balance the observed covariates.

8.2 Inverse-Probability Weighting

Using PSM to estimate ATT has some limitations that may constrain it practical application. Matching algorithms such as Radius or Nearest Neighbor tend to omit a significant proportion of the population size when comparison group is being constructed, thus limiting the ability to provide generalised ATT results. That is why Cassel, et al. (1983), Rosenbaum (1987), and Hirano & Imbens (2001) recommend applying IPW to adjust for this confounding. Compared to PSM the IPW matching requires fewer distributional assumptions about the underlying data[90]. In addition, IPW can incorporate time-dependent covariates and deal with missing observations (Curtis, et al., 2007). Furthermore, Handouyahia, et al. (2013) Finds that for most cases, IPW surpasses Kernel matching in terms of estimate precision. IPW appears to be often superior on technical grounds compared to Kernel matching and offers a strong practical advantage (Handouyahia, et al., 2013).

As discussed in section 5.1 the propensity score is a subject's probability of selection into treatment (IT adoption), conditional on observed covariates. Weighting subjects by the inverse probability of treatment received creates a synthetic sample in which assignment into treatment is independent of measured covariates. IPW using the propensity score therefore allows us to obtain unbiased estimates of average treatment effects (Austin & Stuart, 2015). Considering equation 5.1 that defines the propensity score as the probability of a subject receiving the treatment based on their observed covariates, the inverse probability of treatment weight may be defined as follows (Hirano & Imbens, 2001, p.264)

[90] Thanks to its semi-parametric nature (Hirano & Imbens, 2001).

$$w = \frac{D}{P_r} + \frac{1-D}{1-P_r} \tag{8.1}$$

Where weight of each subject is equal to the subject's inverse probability of receiving the treatment that the subject received[91] (Rosenbaum, 1987). To estimate the ATT, we further have to multiply these weights by P_r (Hirano & Imbens, 2001, p.266)

$$w_{ATT} = D + \frac{P_r(1-D)}{1-P_r} \tag{8.2}$$

Here, the treated group is used as the reference population to which the treated and control groups are standardised. When the propensity score model is correctly specified, both estimators provide consistent estimation of the treatment effect (Lunceford & Davidian, 2004).

A possible problem might represent the fact that using the weights on treated subject with very low propensity score might result in a large weight for the given subject. Similarly, large weight might be also received by a control subject with very high propensity score. This might cause bias and increased variability of the estimates (Cole & Hernan, 2008).

IPW ATT estimates for benchmark model are presented in Table 7. Estimates for adjusted benchmark model (robustness testing) are presented in Table 10. Estimated coefficients have in most cases the same properties and are consistent with those values estimated by PSM methods. In general, however, the IPW estimates tends to be more significant. For example, benchmark estimates of ATT for inflation variability among industrial and developing countries are 3 times more significant than estimates provided by PSM. Unfortunately, the statistical significance is also of similar nature as the PSM's. Only the estimates for inflation variability among pooled and developing countries are found to be highly statistically significant.

Since, the objective of IPW is to create a weighted sample in which the distribution of covariates is similar between treated and control group, we apply similar requirements on proper propensity score specification and its balancing properties that were discussed in sections 7.2 (Austin & Stuart, 2015). I use standardise bias to assess the balancing properties of the model. This allows me to compare the mean terms between treated and control group.

[91] Meaning both inverse probability of remaining control or inverse probability of remaining treated

Table 36 & Table 37 present details on balancing properties for the benchmark model and for the model adjusted for robustness testing respectively.

For the benchmark model the mean standardise bias peaks at 8.2 per cent for pooled dataset and never exceeds the 10 percentage threshold. The matching appears to be even more successful for industrial and developing dataset, for which the mean biases are 2.1 per cent and 3.4 per cent respectively. On the individual basis the biases for industrial and developing datasets never exceeds 10 percentage threshold. However, for the pooled dataset the biases are generally higher and peak at 19.3 per cent for Financial Depth.

The matching quality appears to be less satisfactory for the model adjusted for robustness testing. While the overall mean biases do not exceed 10 percentage threshold for any dataset, inspecting biases on individual basis reveals several drawbacks. Standardise biases exceed the 10 percentage threshold at least once in each of the datasets and peaks at 16.2 per cent. Furthermore, matching for industrial countries failed in two cases in decreasing the bias between treated and control group. This does not have to be interpreted directly as violation of the *conditional independence assumption* however the goal of matching is to reduce this bias which clearly did not happen. This may adversely affect quality of the specific ATT estimate.

Summary

The goal of this thesis was to assess what effect, if any, the IT adoption has had on the relation between output and the inflation, on inflation variability, and on output variability. I implemented several propensity score matching methods to take into account the problems of non-random experiment.

I assigned each of the countries into either developing or industrial dataset, and eliminated observations with extreme values of observed characteristics. I estimated the propensity score for each of the datasets using a probit regression and a variety of outcome-independent variables, including lagged inflation, broad money growth, trade openness, financial depth, and net national income growth. Moreover, I introduced a variety of other control variables to provide evidence that the benchmark model does not omit an important variable and does not suffer from misspecification The matching was subsequently conducted using *Nearest Neighbor, 3 Nearest Neighbor, Radius* matching with various calipers, and *Kernel* matching. Apart from inflation variability, the majority of ATT estimates are statistically insignificant. The results indicate that both industrial and developing IT countries exhibit lower inflation levels and at the same time higher output growth than non-IT countries. Moreover, it appears that both industrial and developing countries achieve combination of lower inflation variability and outcome variability compared to non-IT countries.

To provide evidence on the quality and reliability of the estimates I assessed in detail balancing properties of the models, including assessment of standardise biases, area of common support, and of overlap. Compared to unmatched sample the matching reduced the biases between control and treatment group by around 90 per cent. In overall the standardised biases of the models appear to be reasonable, and quantitatively better than biases presented in the previous research (Ardakani, et al., 2015). The overlap is generous-enough and should provide sufficient support for matching conducted without the no-replacement condition.

I conducted robustness testing by altering some of the countries' assignment into industrial or developing dataset as well as their identification as IT or non-IT country. Compared to the benchmark it appears that adjustments did not have any major impact on the estimates. In some cases, the adjusted model performed better than the benchmark but in other cases it was the other way around. However, assessing standardised biases reveals that the matching was qualitatively less successful than in the case of the benchmark model.

As the last stage of robustness testing I introduced Inverse Probability Weighting method as a substitute to PSM. Estimated coefficients under IPW have in most cases the same properties and are consistent with those values estimated by PSM methods. In general, the IPW estimates tend to be more significant. I assessed the quality and reliability of the estimates by analysing standardise biases. In overall the matching appears to be of a similar quality as the matching conducted with PSM.

While the PSM represents a great improvement to the literature on treatment effect assessment, there are still several drawbacks to its current state of the art methodology. In order to conduct the matching, the original panel dataset had to be transformed into quasi cross-sectional sample in which all of the observations remain, however the time dimension disappears. There is therefore a room for improvement especially in the area of panel data handling.

References

Aizenman, J., Hutchinson, M. & Noy, I., 2011. Inflation targeting and real exchange rates in emerging markets. *World Development.*

Ammer, J. & Freeman, R. T., 1995. Inflation targeting in the 1990s: The experience of New Zealand, Canada, and the United Kingdom. *Journal of Economics and Business 47,* pp. 165-192.

Ardakani, O. M., Kishor, K. N. & Song, S., 2015. On the effectiveness of inflation targeting: evidence from semi/nonparametric approach. *University of Wisconsin-Milwaukee.*

Augurzky, B. & Schmidt, C., 2000. The propensity score: A means to an end. *Working paper, University of Heidelberg.*

Austin, P. C., 2008. A critical appraisal of propensity score matching in the medical literature from 1996 to 2003. *Statistics in Medicine,* pp. 2037-2049.

Austin, P. C., 2009a. The relative ability of different propensity-score methods to balance measured covariates between treated and untreated subjects in observational studies. *Medical Decision Making,* pp. 661-667.

Austin, P. C., 2009b. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine,* pp. 3038-3107.

Austin, P. C., 2011a. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics,* pp. 150-161.

Austin, P. C., 2011b. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research,* pp. 399-424.

Austin, P. C. & Stuart, E. A., 2015. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine.*

Ball, L. et al., 1988. The new keynesian economics and output-inflations trade-off. *Brookings Institution Press,* 1988(1), pp. 1-82.

Ball, L. & Sheridan, N., 2003. Does inflation targeting matter? *University of Chicago Press.*

Bernanke, B. S., Laubach, T., Mishkin, F. S. & Posen, A. S., 1999. Inflation targeting: Lessons from the international experience. *Princeton University Press.*

Bernanke, B. S. & Mishkin, F. S., 1997. Inflation targeting: A new framework for monetary policy? *National Bureau of Economic Research.*

Bryson, A., Dorsett, R. & Purdon, S., 2002. The use of propensity score matching in the evaluation of labour market policies. *Working paper No. 4, Department for Work an Pensions.*

Caliendo, M., Hujer, R. & Thomsen, S., 2005. The employment effects of job creation schemes in Germany - A microeconomic evaluation. *IZA.*

Caliendo, M. & Kopeinig, S., 2005. Some practical guidance for the implementation of propensity score matching. *IZA DP*.

Calvo, G. A. & Mishkin, F. S., 2003. The mirage of exchange rate regimes for emerging countries. *Journal of Economic Perspectives.*

Carare, A. & Stone, M. R., 2006. Inflation targeting regimes. *European Economic Review*, pp. 1297-1373.

Cassel, C., Sarndal, C. & Wretman, J., 1983. Some uses of statistical models in connection with the nonresponse problem. *Academic Press.*

Cave, G. & Bos, H., 1995. The value of a GED in a choice-based experimental sample. *Manpower Demonstration Research Corporation.*

Cecchetti, S. G. & Ehrmann, M., 2002. Does inflation targeting increase output volatility? An international comparison of policymakers' preferences and outcomes. *Monetary Policy: Rules and Transmission Mechanisms*, pp. 247-274.

Clardia, R. H., Gali, J. & Gertler, M., 1998. Monetary policy rules in practice: Some international evidence. *European Economic Review*, pp. 1033-1067.

Cochran, W. G. & Chambers, S. P., 1965. The planning of observational studies of human populations. *Journal of Royal Statistical Society*, pp. 234-266.

Cochran, W. G. & Rubin, D. B., 1973. Controlling bias in observational studies: A review. *The Indian Journal of Statistics*, pp. 417-466.

Cole, S. & Hernan, M. A., 2008. Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology*, pp. 656-664.

Curtis, L. H. et al., 2007. Using inverse probability-weighted estimators in comparative effectiveness analyses with observational databases. *Medical Care*, pp. 103-107.

Czajka, J. L., Hirabayashi, S. M., Little, R. J. A. & Rubin, D. B., 1992. Projecting from advance data using propensity modeling: An application to income and tax statistics. *Journal of Business & Economic Statistics*, pp. 117-131.

de Mendonca, H. F. & de Guimaraes, G. J., 2012. Is inflation targeting a good remedy to control inflation? *Journal of Development Economics*, 98(2), pp. 178-191.

Dehejia, R. H. & Wahba, S., 2002. Propensity score matching methods for nonexperimental causal studies. *Economics and Statistics 84,* pp. 151-161.

Dinardo, J. & Tobias, J., 2001. Nonparametric density and regression estimation. *Journal of Economic Prospective,* pp. 11-28.

Dueker, M. J. & Fischer, A. M., 1996. Inflation targeting in a small open economy: Empirical results from Switzerland. *Journal of Monetary Economics 37(1),* pp. 89-103.

Flury, B. K. & Riedwyl, H., 1986. Standard distance in univariate and multivariate analysis.. *The American Statistician,* pp. 249-251.

Fraga, A., Goldfajn, I. & Minella, A., 2003. Inflation targeting in emerging market economies. *NBER Macro Annual.*

Friedman, M., 1968. The role of monetary policy. *American Economic Review,* pp. 1-17.

FTSE, 2016. [Online] Available at: http://www.ftse.com/products/downloads/FTSE-Country-Classification-Update_latest.pdf

Gertler, M., 2005. Comment on L. Ball and N. Sheridan, "Does inflation targeting matter?". *University of Chicago Press,* pp. 276-281.

Groenevald, H. M., Koedijk, K. G. & Kool, C. J. M., 1998. Inflation dynamics and monetary strategies: Evidence from six industrialized countries. *Open Economies Review,* pp. 21-38.

Groeneveld, J. M., 1998. Inflation patterns and monetary policy: Lessons for the European Central Bank. *Edward Elgar.*

Gu, X. S. & Rosenbaum, P. R., 1993. Comparison of multivariate matching methods: Structures, distances and algorithms. *Journal of Computational and Graphical Statistics 2,* pp. 405-420.

Hahn, E. D. & Soyer, R., 2005. Probit and logit models: Differences in the multivariate realm.

Hamilton, K. & Ley, E., 2010. Measuring national income and growth in resource-rich. *Economic Premise.*

Hammond, G., 2012. State of the art of inflation targeting. *Handbook No. 29, Bank of England.*

Hammond, G., Kanbur, R. & Prasad, E. S., 2009. Monetary policy challenges for emerging market economies. *Global Economy & Development.*

Handouyahia, A., Haddad, T. & Eaton, F., 2013. Kernel matching versus inverse probability weighting: A comparative study. *International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering,* pp. 1218-1233.

Heckman, J. J., Ichimura, H. & Todd, P., 1998. Characterizing selection bias using experimental data. *Econometrica 66,* pp. 1017-1098.

Heckman, J. J., Ichimura, H. & Todd, P., 1997. Matching as an econmetric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies,* pp. 605-654.

Heckman, J. J., LaLonde, R. & Smith, J., 1999. The economics and econometrics of active labor market programs. *Handbook of Labor Economics Vol III,* pp. 1865-2097.

Heenan, G., Marcel, P. & Roger, S., 2006. Implementing inflation targeting: Institutional arrangements, target design, and communication. *International Monetary Funds.*

Hirano, K. & Imbens, G., 2001. Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes Research Methodology,* pp. 259-278.

Ho, D. E., Imai, K., King, G. & Stuart, E., 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference.. *Political Analysis,* pp. 199-236.

Chih-Lin, L., 2011. Propensity score matching in observational studies with multiple time points. *Disertation, Ohio State University.*

Imai, K., King, G. & Stuart, E. A., 2008. Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society,* pp. 481-501.

Imbens, G. W., 2004. On parametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics,* pp. 4-29.

IMF, 2005. Does inflation targeting work in emerging markets? *World Economic Outlook,* pp. 161-186.

IMF, 2016. World economic outlook.

J.P. Morgan, 2016. *Emerging markets bond index monitor.* [Online] Available at: https://www.jpmorgan.com/country/US/EN/jpmorgan/investbk/solutions/research/indices/product

Johnson, D. R., 2002. The effect of inflation targeting on the behavior of expected inflation: Evidence from an 11 country panel. *Journal of Monetary Economics 49(8),* pp. 1521-1538.

Kuttner, K. N., 2004. Snapshot of inflation targeting in its adolescence. *Reserve Bank of Australia.*

Kuttner, K. N. & Posen, A. S., 1999. Does talk matter after all? Inflation targeting and central bank behaviour. *Federal Reserve Bank of New York.*

Lechner, M., 1999. Earnings and employment effects of continuous off-the-job training in East Germany after unification. *Journal of Business Economic Statistics,* pp. 74-90.

Leuven, E. & Sianesi, B., 2003. *Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.* [Online] Available at: http://ideas.repec.org/c/boc/bocode/s432001.html

Levin, A., Natalucci, F. M. & Piger, J., 2004. The macroeconomic effects of inflation targeting. *Federal Bank of St. Louis Review 86,* pp. 51-80.

Levin, A. & Piger, J., 2004. Is inflation persistance intristic in industrial economies? *European Central Bank.*

Li, M., 2012. Using the propensity score method to estimate causal effects: A review and practical guide. *Organizational Research Methods,* pp. 1-39.

Lin, S., 2010. On the international effects of inflation targeting. *The Review of Economics and Statistics,* pp. 195-199.

Lin, S. & Ye, H., 2007. Does inflation targeting really make a difference? Evaluating the treatment effect of inflation targeting in seven industrial countries. *Journal of Monetary Economics.*

Lin, S. & Ye, H., 2009. Does inflation targeting make a difference in developing countries?.*Journal of Development Economics 89,* pp. 118-123.

Lin, S. & Ye, H., 2012. What to target? Inflation or exchange rate. *Souther Economic Association,* pp. 1202-1221.

Louviere, J. J., Hensher, D. & Swait, J. D., 2000. Stated choice methods: Analysis and application. *Cambridge University Press.*

Lucas, R. E., 1973. International evidence on output-inflation tradeoffs. *The American Economic Review,* 63(3), pp. 326-334.

Lucotte, Y., 2012. Adoption of inflation targeting and tax revenue performance in emerging market economies: An empirical investigation. *Economic Systems, 36(4),* pp. 609-628.

Lunceford, J. K. & Davidian, M., 2004. Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Statistics in Medicine,* pp. 2937-2960.

Masson, P., Savastano, M. & Sharma, S., 1997. The scope for inflation targeting in developing countries. *IMF,* Issue 97/130.

Miao, Y., 2009. In search of successful inflation targeting: Evidence from an inflation targeting index. *International Monetary Fund.*

Minea, A. & Tapsoba, R., 2014. Does inflation targeting improve fiscal discipline? *Journal of International Money and Finance,* 40(C), pp. 185-203.

Mishkin, F. S., 1999. International experiences with different monetary policy regimes. *Journal of Monetary Economics 43,* pp. 579-605.

Mishkin, F. S., 2004. Can inflation targeting work in emerging market countries? *National Bureau of Economic Research.*

Mishkin, F. S. & Posen, A. S., 1997. Inflation targeting: Lessons from four countries. *National Bureau of Economic Research.*

Mishkin, F. S. & Schmidt-Hebbel, K., 2001. One decade of inflation targeting in the world: What do we know and what do we need to know? *National Bureau of Economic Research.*

Mishkin, F. S. & Schmidt-Hebbel, K., 2007. Does inflation targeting make a difference? *Monetary Policy under Inflation Targeting,* pp. 291-372.

Morgan, S. & Todd, J. L., 2008. A diagnostic routine for the detection of consequential heterogeneity of causal effects. *Sociological Methodology,* pp. 231-281.

Neumann, M. J. M. & von Hagen, J., 2002. Does inflation targeting matter? *Federal Reserve Bank of St. Louis,* pp. 127-148.

Normand, S. L. et al., 2001. Validating recommendations for coronary angiography following an acute myocardial infarction in the lderly: A matched analysis using propensity scores. *Journal of Clinical Epidemiology,* pp. 387-398.

Pétursson, T. G., 2004. The effects of inflation targeting on macroeconomic performance. *Central Bank of Iceland,* p. Working Paper No. 23.

Phelps, E., 1967. Phillips curve, expectations of inflation and optimal unemployment over time. *Economica,* 135(34), pp. 254-281.

Phelps, E., 1972. Inflation policy and unemployment theory. *W. W. Norton.*

Phillips, A. W., 1958. The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861-1957. *Economica,* pp. 283-299.

Romer, D., 1993. Openness and inflation: Theory and evidence. *Quaterly Journal of Economics,* pp. 869-903.

Rosenbaum, P. R., 1987. Model-based direct adjustment. *Journal of the American Statistical Association,* pp. 397-394.

Rosenbaum, P. R., 1995. Observational Studies. *Springer Series in Statistics.*

Rosenbaum, P. R. & Rubin, D. B., 1983. The central role of the propensity score for causal effects. *Biometrika 70,* pp. 41-55.

Rosenbaum, P. R. & Rubin, D. B., 1985. Constructing a control group using multivariate matched sampling methods that Incorporate the propensity score. *American Statistician 39,* pp. 33-38.

Rubin, D. B., 1973. Matching to remove bias in observational studies. *Biometrics 29,* pp. 159-184.

Rubin, D. B. & Thomas, N., 1992. Affinely invariant matching methods with ellipsoidal distributions. *The Annals of Statistics,* pp. 1079-1093.

Samuelson, P. A. & Solow, R. M., 1960. Analytical aspects of anti-inflation policy. *American Economic Review Papers and Proceedings 50(2),* pp. 177-194.

Sianesi, B., 2004. An evaluation of the active labour market programmes in Sweden. *The Review of economics and Statistics,* pp. 133-155.

Smith, J. & Todd, P., 2005. Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics, 125(1-2),* pp. 305-353.

Svensson, L. E., 1996. Inflation forecast targeting: Implementing and monitoring inflation targets. *European Economic Review,* pp. 1111-1146.

Svensson, L. E., 1999. Inflation targeting as a monetary policy rule. *Journal of Monetary Economics,* pp. 607-654.

Taylor, J. B., 1979. Estimation and control of a macroeconomic model with rational expectations. *Econometrica 47,* pp. 1267-1286.

Taylor, J. B., 1993. Discretion versus policy rules in practice. *Stanford University,* pp. 195-214.

Taylor, J. B., 1994. The inflation/output variability trade-off revisited. *Federal Reserve Bank of Boston,* pp. 21-24.

Taylor, J. B., 1996. How should monetary policy respond to shocks while maintaining long-run price stability? - Conceptual issues. *Federal Reserve Bank of Kansas City,* pp. 181-195.

Truman, E. M., 2003. Inflation targeting in the world economy. *Institute for International Economics.*

Vega, M. & Winkelried, D., 2005. Inflation targeting and inflation behaviour: A successful story? *International Journal of Central Banking,* pp. 153-175.

Walker, S. H. & Duncan, D. B., 1967. Estimation of the probability of an event as a function of several independent variables. *Biometrika,* Volume 54, pp. 167-179.

Walsh, C. E., 1998. The new output-inflation trade-off. *Federal Reserve Bank of San Francisco.*

Woodford, M., 1995. Price level determinacy without control of a monetary aggregate. *Carnegie-Rochester Conference Series on Public Policy 43,* pp. 1-46.

Woodford, M., 2004. Inflation targeting and optimal monetary policy. *Federal Reserve Bank of St. Louis,* pp. 15-42.

Wu, T. Y., 2004. Does inflation targeting reduce inflation? An analysis for the OECD industrial countries. *Working Paper Series* .

Appendix

Table 11 - List of developing countries

Benchmark model	Robustness check	EMBI	FTSE	IMF	Benchmark model	Robustness check	EMBI	FTSE	IMF
Argentina	✓	✓		✓	Lebanon	✓	✓		✓
Armenia	✓	✓		✓	Liberia	✓			✓
Azerbaijan	✓	✓		✓	Libya	✓			✓
Bahrain	✓		✓	✓	Lithuania	✓	✓	✓	✓
Bangladesh	✓		✓	✓	Malaysia	✓	✓	✓	✓
Belarus	✓	✓		✓	Mexico	✓	✓	✓	✓
Bolivia	✓	✓		✓	Moldova	✓			✓
Bosnia and Herzegovina	✓			✓	Montenegro	✓			✓
Brazil	✓	✓	✓	✓	Paraguay	✓	✓		✓
Bulgaria	✓		✓	✓	Peru	✓	✓	✓	✓
Colombia	✓	✓	✓	✓	Poland	✓	✓		✓
Costa Rica	✓	✓		✓	Puerto Rico	✓			✓
Croatia	✓	✓	✓	✓	Qatar	✓		✓	✓
Czech Rep	X[92]		✓		Romania	✓	✓	✓	✓
Egypt, Arab Rep.	✓	✓	✓	✓	Saudi Arabia	✓			✓
Estonia	✓		✓	✓	Serbia	✓	✓	✓	✓
Georgia	✓			✓	Slovak Republic	✓	✓	✓	✓
Hungary	✓	✓	✓	✓	Slovenia	X[93]		✓	
Chile	✓	✓	✓	✓	South Africa	✓	✓		✓
China	✓	✓	✓	✓	Thailand	✓		✓	✓
India	✓	✓	✓	✓	Tunisia	✓	✓		✓
Indonesia	✓	✓	✓	✓	Turkey	✓	✓	✓	✓
Israel	X[94]	✓			Ukraine	✓	✓		✓
Jamaica	✓	✓		✓	United Arab Emirates	✓	✓	✓	✓
Kazakhstan	✓	✓		✓	Uruguay	✓	✓		✓
Korea Rep.	X[95]	✓			Vietnam	✓	✓	✓	✓
Latvia	✓	✓	✓	✓					

Source: FTSE (2016), J.P. Morgan (2016), IMF (2016). Countries assigned as IT are listed in column Benchmark model. Sign "✓" means that the country is included in the given list, sign "X" means that the country was excluded from the sample for robustness testing.

[92] For periods 2008-2015

[93] For periods 2003-2015

[94] For periods 1990-2015

[95] For periods 2007-2015

Table 12 - IT countries & adoption dates

	Benchmark model	**Robustness check**	Rose (2007)		Miao (2009)	Hammond (2012)
IT country	Adoption	Adoption	Adoption	Conservative adoption[96]	Adoption	Adoption
Armenia	2006	N/A	N/A	N/A	N/A	2006
Australia	1993	1993	1993	1994	1993	1993
Brazil	1999	1999	1999	1999	1999	1999
Canada	1991	1991	1991	1992	1991	1991
Colombia	1999	1999	1999	1999	1999	1999
Czech Republic	1998	1998	1998	1998	1998	December 1997
Hungary	2001	2001	2001	2001	2001	2001
Chile	1999	1999	1991	1999	1999	1999
Iceland	2001	2001	2001	2001	2001	2001
Indonesia	2005	N/A	N/A	N/A	N/A	2005
Israel	1992	1997	1992	1997	1997	1992
Korea Republic	1998	1998	1998	1998	2001	1998
Mexico	2001	2001	1999	2001	2001	2001
New Zealand	1990	1990	1990	1990	1990	December 1989
Norway	2001	2001	2001	2001	2001	2001
Peru	2002	2002	2002	2002	2002	2002
Poland	1998	1998	1998	1998	1998	1998
Romania	2005	N/A	N/A	N/A	N/A	2005
Serbia	2012	N/A	N/A	N/A	N/A	2012
South Africa	2000	2000	2000	2000	2000	2000
Sweden	1993	1995	1993	1995	1993	1993
Switzerland	2000	N/A	2000	2000	N/A	N/A
Thailand	2000	2000	2000	2000	2000	2000
Turkey	2006	2006	N/A	N/A	N/A	2006
United Kingdom	1992	1992	1992	1992	1992	1992

Source: Rose (2007), Miao (2009), Hammond (2012). Column Benchmark model specifies adoption years as accounted for in the benchmark model. Column Robustness check assigns reclassified adoption year for purpose of robustness testing in section 7.1

[96] Rose (2007) for some countries provides also alternative, “conservative” adoption dates.

Table 13 - Descriptive Statistics, Developing countries

Variable	Obs	Mean	Std. Dev.	Min	Max
CP_infl_VAR	1202	.112306	1.282397	.0000125	.8998939
log_CPI	1202	.0307221	.0301874	-.0447821	.1723819
GDP_pc_VAR	1202	.0205738	.0263743	6.52e-06	.4639181
RpcGDPGr	1202	.0329777	.0437603	-.2391497	.293171
FinD	1186	.4655505	.3242717	.0018587	1.665041
CP_infl_1	1202	.1098099	.2745671	-.0979765	.2976935
NNIpcGr	1202	.0353893	.0640449	-.2362766	.2753107
Trade	1202	.8334415	.4229667	.1200868	3.216317
BoP	1156	-.0254466	.0857236	-.8005137	.4222732
Esize	1202	.0047474	.0109913	.0000109	.1478381
BMGr	1173	.1774412	.1576963	-.479132	1.543908
FisBal	788	-.0163638	.0340944	-.1502366	.1956639
ResR	1190	.1736897	.1743435	.0028399	2.527196
PopGr	1198	.0110773	.0171696	-.0285097	.1762477

Table 14 - Descriptive Statistics – Developing & Inflation targeting countries

Variable	Obs	Mean	Std. Dev.	Min	Max
CP_infl_VAR	247	.0127116	.0109218	.0000605	.0513161
log_CPI	247	.0193993	.01213	-.0043266	.0596214
GDP_pc_VAR	247	.0165638	.0182565	.0000141	.1647922
RpcGDPGr	247	.0262197	.0316626	-.1389178	.1428382
FinD	247	.6129117	.3787528	.0883746	1.601249
CP_infl_1	247	.049458	.0328843	-.0084572	.1900595
NNIpcGr	247	0.031722	.0478628	-.1968372	.0180134
Trade	247	.7424502	.3405173	.2098214	1.712419
BoP	245	-.0205715	.0438351	-.2138715	.1063979
Esize	247	.0070506	.007221	.0001251	.0359055
BMGr	247	.127498	.1147735	-.479132	.8520308
FisBal	208	-.021614	.0298315	-.1028541	.076946
ResR	247	.1858177	.0875238	.0370629	.4914014
PopGr	247	.0084907	.008424	-.0166638	.0345547

Table 15 - Descriptive Statistics - Developing & Non-targeting countries

Variable	Obs	Mean	Std. Dev.	Min	Max
CP_infl_VAR	955	.138065	1.43773	.0000125	.8998939
log_CPI	955	.0336506	.0326726	-.0447821	.1723819
GDP_pc_VAR	955	.0216109	.0280092	6.52e-06	.4639181
RpcGDPGr	955	.0347256	.0462317	-.2391497	.293171
FinD	939	.4267878	.2966427	.0018587	1.665041
CP_infl_1	955	.1254192	.3185211	-.0979765	.2976935
NNIpcGr	955	.0366336	.0686613	-.2362766	.2753107
Trade	955	.8569754	.4388877	.1200868	3.216317
BoP	911	-.0267577	.0938278	-.8005137	.4222732
Esize	955	.0041517	.011701	.0000109	.1478381
BMGr	926	.1907629	.1648078	-.4547297	1.543908
FisBal	580	-.0144809	.0353328	-.1502366	.1956639
ResR	943	.170513	.1905689	.0028399	2.527196
PopGr	951	.0117491	.0187316	-.0285097	.1762477

Table 16 - Descriptive Statistics - Industrial countries

Variable	Obs	Mean	Std. Dev.	Min	Max
CP_infl_VAR	766	.0099645	.0171899	9.09e-06	.2563038
log_CPI	766	.0130748	.0157507	-.1537413	.1195437
GDP_pc_VAR	766	.012887	.0122373	2.08e-06	.1115893
RpcGDPGr	766	.0176058	.0256946	-.0899796	.1321649
FinD	722	1.005053	.4792151	.2188235	3.121536
CP_infl_1	766	.0346765	.0496835	-.2981268	.2917897
NNIpcGr	766	.0192264	.0332873	-.115752	.2215677
Trade	766	.8299687	.6611319	.1592399	4.396567
BoP	702	.0056019	.0609642	-.2366981	.2610381
Esize	766	.0300217	.0577386	.0001891	.3429952
BMGr	707	.0797331	.1079198	-.2862978	1.25031
FisBal	711	-.020817	.0502268	-.3236632	.2033753
ResR	766	.1033261	.1724569	.0009699	.9900246
PopGr	675	1.022738	.0607663	.7108704	1.285791

Table 17 - Descriptive Statistics - Industrial & Inflation targeting countries

Variable	Obs	Mean	Std. Dev.	Min	Max
CP_infl_VAR	167	.0077236	.0077433	.0000885	.0463227
log_CPI	167	.0094439	.0076371	-.0049966	.0518399
GDP_pc_VAR	167	.0114488	.0114674	.0001588	.0583202
RpcGDPGr	167	.0145702	.0209359	-.0598896	.0727438
FinD	147	1.194373	.4319339	.3168045	3.121536
CP_infl_1	167	.0236254	.0189049	-.0069254	.1267819
NNIpcGr	167	.0178731	.0319304	-.0981582	.0806120
Trade	167	.680824	.2044366	.3547598	1.324944
BoP	167	.0062024	.0706936	-.2366981	.1618676
Esize	167	.0146576	.0141576	.0001909	.0527827
BMGr	167	.0786529	.1239044	-.2862978	1.25031
FisBal	165	-.0033886	.0521289	-.1343546	.1890385
ResR	167	.0987781	.1302672	.0157251	.8529074
PopGr	147	1.019195	.0382166	.9571235	1.209008

Table 18 - Descriptive Statistics - Industrial & Non-targeting countries

Variable	Obs	Mean	Std. Dev.	Min	Max
CP_infl_VAR	599	.0105892	.0189624	9.09e-06	.2563038
log_CPI	599	.0140871	.0172182	-.1537413	.1195437
GDP_pc_VAR	599	.013288	.012423	2.08e-06	.1115893
RpcGDPGr	575	.01754914	.0250539	-.0899812	.1321649
FinD	599	.0377575	.0549088	-.2981268	.8422198
CP_infl_1	599	.87155	.7345848	.1592399	.2917897
NNIpcGr	599	.0196242	.0336955	-.115752	.2215677
Trade	535	.0054145	.0576649	-.1536005	.2610381
BoP	599	.0343053	.0642247	.0001891	.3429952
Esize	599	.0343053	.0642247	.0001891	.3429952
BMGr	540	.0800672	.1025991	-.1723939	1.089361
FisBal	546	-.0260838	.0484668	-.3236632	.2033753
ResR	599	.1045941	.1825626	.0009699	.9900246
PopGr	528	1.023725	.0656768	.7108704	1.285791

Table 19 - Descriptive Statistics - Pooled dataset

Variable	Obs	Mean	Std. Dev.	Min	Max
CP_infl_VAR	1968	.0724719	1.003355	9.09e-06	.8998939
log_CPI	1968	.0238533	.0269622	-.1537413	.1723819
GDP_pc_VAR	1968	.0175819	.0222938	2.08e-06	.4639181
RpcGDPGr	1968	.0269945	.0384993	-.2391497	.293171
FinD	1908	.6697018	.4697493	.0018587	3.121536
CP_infl_1	1968	.0805659	.2965917	-.2981268	.2976935
NNIpcGr	1968	.0284977	.0537372	-.2362766	.2753107
Trade	1968	.8320898	.5284213	.1200868	4.396567
BoP	1858	-.0137157	.078742	-.8005137	.4222732
Esize	1968	.0145849	.0390161	.0000109	.3429952
BMGr	1489	.6412992	.4006945	.0682303	2.620777
FisBal	1499	-.018476	.0425598	-.3236632	.2033753
ResR	1956	.1461342	.1769302	.0009699	2.527196
PopGr	1963	.009494	.0142167	-.0285097	.1762477

Table 20 - Descriptive Statistics - Pooled dataset, Inflation targeting countries

Variable	Obs	Mean	Std. Dev.	Min	Max
CP_infl_VAR	414	.0106996	.0100575	.0000605	.0513161
log_CPI	414	.0153834	.0116187	-.0049966	.0596214
GDP_pc_VAR	414	.0145005	.0160529	.0000141	.1647922
RpcGDPGr	414	.0215205	.0283912	-.1389178	.1428382
FinD	394	.8298527	.4882477	.0883746	3.121536
CP_infl_1	414	.0390376	.0308018	-.0084572	.1900595
NNIpcGr	414	.0261824	.0427135	-.1968372	.1801342
Trade	414	.7175913	.2945858	.2098214	1.712419
BoP	412	-.009719	.0577274	-.2366981	.1618676
Esize	414	.0101191	.0112063	.0001251	.0527827
BMGr	394	.7470272	.3580246	.1828698	1.894501
FisBal	373	-.0135518	.042135	-.1343546	.1890385
ResR	414	.1507075	.11494	.0157251	.8529074
PopGr	413	.008797	.0071012	-.0166638	.0345547

Table 21 - Descriptive Statistics - Pooled dataset, Non-targeting countries

Variable	Obs	Mean	Std. Dev.	Min	Max
CP_infl_VAR	1554	.0889286	1.128618	9.09e-06	.8998939
log_CPI	1554	.0261098	.0293366	-.1537413	.1723819
GDP_pc_VAR	1554	.0184028	.0236172	2.08e-06	.4639181
RpcGDPGr	1554	.0284529	.0406552	-.2391497	.293171
FinD	1514	.6280245	.4558397	.0018587	2.535634
CP_infl_1	1554	.0916294	.3325392	-.2981268	.2976935
NNIpcGr	1554	.0292376	.0568072	-.2362766	.2753107
Trade	1554	.8625933	.571106	.1200868	4.396567
BoP	1446	-.0148544	.0837523	-.8005137	.4222732
Esize	1554	.0157746	.0434503	.0000109	.3429952
BMGr	1095	.6032564	.408483	.0682303	2.620777
FisBal	1126	-.0201072	.0425927	-.3236632	.2033753
ResR	1542	.1449064	.1901751	.0009699	2.527196
PopGr	1550	.0096798	.01557	-.0285097	.1762477

Table 22 - Standardise Bias, Nearest Neighbor

			Pooled				Developing				Industrial			
		Unmatched	Mean				Mean				Mean			
Matching	Variable	Matched	Treated	Control	%bias	%reduct bias	Treated	Control	%bias	%reduct bias	Treated	Control	%bias	%reduct bias
	NNIpcGr	U	.02237	.02868	-17.9		.02622	.03544	-23.4		.0159	.01697	-4.4	
		M	.02237	.0219	4.9	71.2	.02622	.02806	-4.7	80.0	.0159	0.1524	3.8	13.6
	FinD	U	.04005	.09321	-22.0		.61291	.42587	54.8		-.01037	-.0264	34.1	
		M	.04005	.0383	0.7	96.7	.61291	.56563	13.1	74.7	-.01037	-.1843	18.2	46.6
Nearest	CPI_infl_1	U	.72132	.86086	-30.2		.04946	.12529	-25.3		.02424	.03761	-32.4	
Neighbor		M	.72132	.69586	6.7	74.6	.04946	.0445	1.7	93.5	.02424	.02386	0.9	97.2
	Trade	U	.1107	.15041	-28.5		.74245	.84906	-27.2		.6858	.8813	-34.7	
		M	.1107	.10867	7.6	69.7	.74245	.73134	2.8	89.6	.6858	.66038	4.5	87.0
	BMGr	U	.82985	.62644	42.8		.1275	.19067	-44.6		.08247	.08064	1.6	
		M	.82985	.81314	6.6	77.6	.1275	.12938	-1.3	97.0	.08247	.08410	-1.5	6.25

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.072	0.000	33.3*	0.02	0.135	0.000	35.1	27.2	0.079	0.000	68.4*	0.51
Matched	0.017	0.502	5.3	0.01	0.007	0.402	4.9	2.8	0.037	0.089	5.8	0.12

Table 23 - Standardise Bias, 3 Nearest Neighbor

			Pooled				Developing				Industrial			
		Unmatched	Mean				Mean				Mean			
Matching	Variable	Matched	Treated	Control	%bias	%reduct bias	Treated	Control	%bias	%reduct bias	Treated	Control	%bias	%reduct bias
	NNIpcGr	U	.02237	.02868	-17.9		.02622	.03544	-23.4		.0159	.01697	-4.4	
		M	.02237	.02051	5.3	70.6	.02622	.02491	3.3	85.8	.0159	0.1574	1.2	72.7
	FinD	U	.04005	.09321	-22.0		.61291	.42587	54.8		-.01037	-.0264	34.1	
		M	.04005	.03597	1.7	92.3	.61291	.58632	7.8	85.8	-.01037	-.01613	15.8	53.7
3 Nearest	CPI_infl_1	U	.72132	.86086	-30.2		.04946	.12529	-25.3		.02424	.03761	-32.4	
Neighbors		M	.72132	.71991	0.3	99.0	.04946	.04618	1.1	95.7	.02424	.02596	-4.2	87.1
	Trade	U	.1107	.15041	-28.5		.74245	.84906	-27.2		.6858	.8813	-34.7	
		M	.1107	.1033	5.3	81.4	.74245	.73678	1.4	94.7	.6858	.67171	2.5	92.8
	BMGr	U	.82985	.62644	42.8		.1275	.19067	-44.6		.08247	.08064	1.6	
		M	.82985	.80109	9.9	76.9	.1275	.12967	-1.5	96.6	.08247	.08197	0.5	68.8

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.072	0.000	33.3*	0.02	0.135	0.000	35.1	27.2	0.079	0.000	68.4*	0.51
Matched	0.013	0.771	4.5	0.01	0.003	0.811	3.0	1.5	0.036	0.118	5.2	0.11

Table 24 - Standardise Bias, Radius 0.03

			Pooled				**Developing**				**Industrial**			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	|bias|	Treated	Control	%bias	bias
	NNIpcGr	U	.02237	.02868	-17.9		.02622	.03544	-23.4		.0159	.01697	-4.4	
		M	.02237	.02068	4.8	73.3	.02622	.0272	-2.5	89.4	.01589	0.1565	1.2	72.6
	FinD	U	.04005	.09321	-22.0		.61291	.42587	54.8		-.01037	-.0264	34.1	
		M	.04005	.03671	1.4	93.7	.61291	.58889	7.0	87.2	-.01076	-.01547	15.2	55.4
Radius	CPI_infl_1	U	.72132	.86086	-30.2		.04946	.12529	-25.3		.02424	.03761	-32.4	
(0.03)		M	.72132	.72188	-0.1	99.6	.04946	.04704	0.8	96.8	.02413	.0256	-3.6	89.0
	Trade	U	.1107	.15041	-28.5		.74245	.84906	-27.2		.6858	.8813	-34.7	
		M	.1107	.10453	4.4	84.5	.74245	.73677	1.5	94.7	.68506	.67618	1.6	95.5
	BMGr	U	.82985	.62644	42.8		.1275	.19067	-44.6		.08247	.08064	1.6	
		M	.82985	.80914	8.1	81.1	.1275	.12803	-0.4	99.2	.07892	.07920	0.8	47.9

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.072	0.000	33.3	0.02	0.135	0.000	35.1	27.2	0.079	0.000	68.4	0.51
Matched	0.012	0.814	3.8	0.01	0.002	0.947	2.4	1.5	0.033	0.143	4.5	0.11

Table 25 - Standardise Bias, Radius 0.01

			Pooled				**Developing**				**Industrial**			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	|bias|	Treated	Control	%bias	bias
	NNIpcGr	U	.02237	.02868	-17.9		.02622	.03544	-23.4		.0159	.01697	-4.4	
		M	.02237	.02032	5.8	67.5	.02641	.02605	0.9	96.1	.01595	0.1531	1.3	72.6
	FinD	U	.04005	.09321	-22.0		.61291	.42587	54.8		-.01037	-.0264	34.1	
		M	.04005	.03617	1.6	92.7	.59731	.57957	5.2	90.5	-.01094	-.01517	15.0	56.0
Radius	CPI_infl_1	U	.72132	.86086	-30.2		.04946	.12529	-25.3		.02424	.03761	-32.4	
(0.01)		M	.72132	.72178	-0.1	99.7	.04933	.0475	0.6	97.6	.02411	.02532	-2.9	91.0
	Trade	U	.1107	.15041	-28.5		.74245	.84906	-27.2		.6858	.8813	-34.7	
		M	.1107	.10342	5.2	81.7	.7458	.73931	1.7	93.9	.68438	.68554	-0.2	99.4
	BMGr	U	.82985	.62644	42.8		.1275	.19067	-44.6		.08247	.08064	1.6	
		M	.82985	.80759	8.6	79.9	.12811	.12886	-0.5	98.8	.07084	.06952	1.1	27.7

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.072	0.000	33.3	0.02	0.135	0.000	35.1	27.2	0.079	0.000	68.4	0.51
Matched	0.012	0.792	4.3	0.01	0.001	0.981	1.8	0.9	0.030	0.167	4.1	0.11

Table 26 - Standardise Bias, Radius 0.005

			Pooled				Developing				Industrial			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	|bias|	Treated	Control	%bias	bias
	NNIpcGr	U	.02237	.02868	-17.9		.02622	.03544	-23.4		.0159	.01697	-4.4	
		M	.02237	.0199	7.0	60.9	.02675	.02673	0.0	99.8	.01595	.01542	1.1	74.9
	FinD	U	.04005	.09321	-22.0		.61291	.42587	54.8		-.01037	-.0264	34.1	
		M	.04005	.03648	1.5	93.3	.58245	.5661	4.8	91.3	-.01094	-.01488	14.8	56.6
Radius	CPI_infl_1	U	.72132	.86086	-30.2		.04946	.12529	-25.3		.02424	.03761	-32.4	
(0.005)		M	.72132	.72026	0.2	99.2	.04935	.04837	0.3	98.7	.02411	.02549	-3.3	89.7
	Trade	U	.1107	.15041	-28.5		.74245	.84906	-27.2		.6858	.8813	-34.7	
		M	.1107	.10296	5.6	80.5	.745	.73186	3.4	87.7	.68438	.69361	-1.6	95.3
	BMGr	U	.82985	.62644	42.8		.1275	.19067	-44.6		.08247	.08064	1.6	
		M	.82985	.80984	8.0	81.3	.12867	.13052	-1.3	97.1	.07084	.07152	-1.5	6.25

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.072	0.000	33.3	0.02	0.135	0.000	35.1	27.2	0.079	0.000	68.4	0.51
Matched	0.012	0.768	4.5	0.01	0.001	0.991	2.0	1.3	0.037	0.174	4.46	0.11

Table 27 - Standardise Bias, Radius 0.001

			Pooled				Developing				Industrial			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	|bias|	Treated	Control	%bias	bias
	NNIpcGr	U	.02237	.02868	-17.9		.02622	.03544	-23.4		.0159	.01697	-4.4	
		M	.02265	.02015	6.9	61.5	.02776	.02414	9.2	60.8	.01524	0.1392	1.9	56.7
	FinD	U	.04005	.09321	-22.0		.61291	.42587	54.8		-.01037	-.0264	34.1	
		M	.04032	.03648	1.6	92.8	.50174	.48943	3.6	93.4	-.01052	-.01634	17.9	47.5
Radius	CPI_infl_1	U	.72132	.86086	-30.2		.04946	.12529	-25.3		.02424	.03761	-32.4	
(0.001)		M	.72407	.72765	-0.8	97.4	.05154	.04986	0.6	97.8	.02308	.02631	-7.8	75.9
	Trade	U	.1107	.15041	-28.5		.74245	.84906	-27.2		.6858	.8813	-34.7	
		M	.11172	.10299	6.3	78.0	.76364	.77428	-2.7	90.0	.70278	.67983	4.1	88.3
	BMGr	U	.82985	.62644	42.8		.1275	.19067	-44.6		.08247	.08064	1.6	
		M	.82025	.81003	7.9	81.5	.13615	.13427	1.3	97.0	.06565	.06524	0.3	77.6

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.072	0.000	33.3	0.02	0.135	0.000	35.1	27.2	0.079	0.000	68.4	0.51
Matched	0.014	0.754	4.7	0.01	0.003	0.927	3.2	2.7	0.039	0.105	4.4	0.12

Table 28 - Standardise Bias, Kernel Matching

			Pooled				Developing				Industrial			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	|bias|	Treated	Control	%bias	bias
	NNIpcGr	U	.02237	.02868	-17.9		.02622	.03544	-23.4		.0159	.01697	-4.4	
		M	.02237	.01946	8.2	53.9	.02622	.02721	-2.5	89.3	.0159	.01378	2.0	54.4
	FinD	U	.04005	.09321	-22.0		.61291	.42587	54.8		-.01037	-.0264	34.1	
		M	.04005	.03794	0.9	96.0	.61291	.58527	8.1	85.2	-.01037	-.01319	13.7	59.8
Kernel Matching	CPI_infl_1	U	.72132	.86086	-30.2		.04946	.12529	-25.3		.02424	.03761	-32.4	
		M	.72132	.69791	5.1	83.2	.04946	.04281	2.2	91.2	.02424	.02434	-0.2	99.3
	Trade	U	.1107	.15041	-28.5		.74245	.84906	-27.2		.6858	.8813	-34.7	
		M	.1107	.10345	6.2	79.2	.74245	.75183	-2.4	91.2	.6858	.68806	-0.4	98.8
	BMGr	U	.82985	.62644	42.8		.1275	.19067	-44.6		.08247	.08064	1.6	
		M	.82985	.81129	7.7	82.0	.1275	.12845	-0.7	98.5	.08247	.08342	-1.5	6.25
		Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
		Unmatched	0.072	28.5	33.3	0.02	0.135	0.000	35.1	27.2	0.079	0.000	68.4	0.51
		Matched	0.015	7.4	4.9	0.01	0.007	0.400	3.2	2.4	0.040	.0138	3.6	0.12

Figure 4 - Standardise Bias, Nearest Neighbor, Developing countries

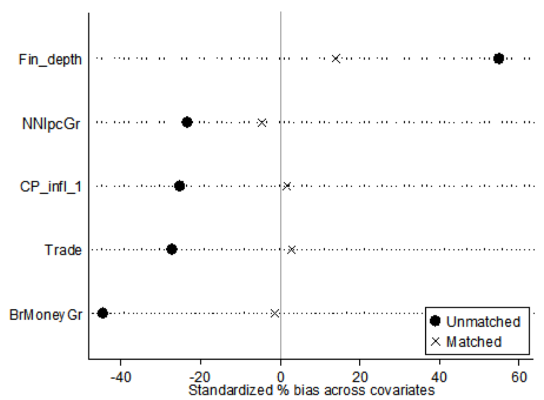


Figure 5 - Standardise Bias, 3 Nearest Neighbor, Developing countries

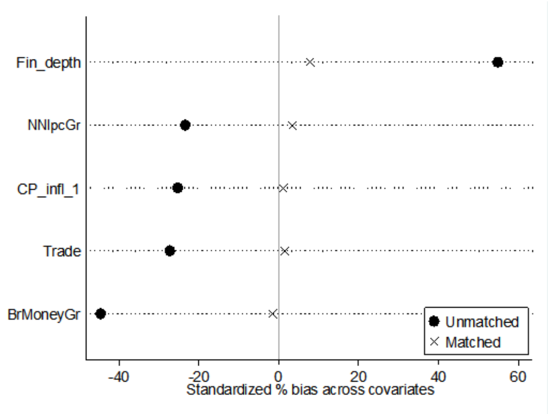


Figure 6 - Standardise Bias, Radius 0.03, Developing countries

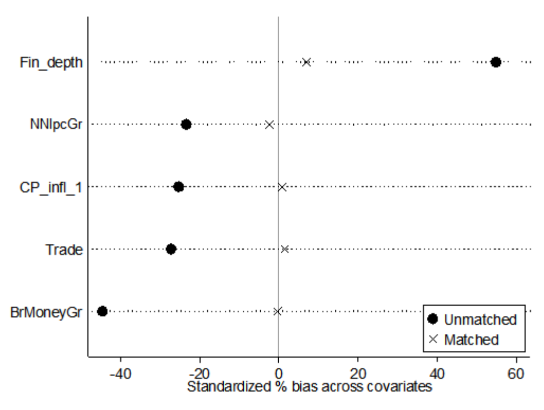


Figure 7 - Standardise Bias, Radius 0.01, Developing countries

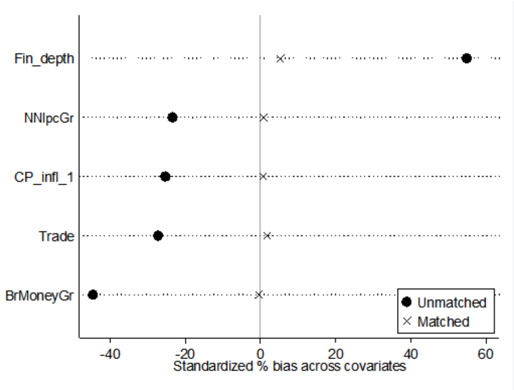


Figure 8 - Standardise Bias, Radius 0.005, Developing countries

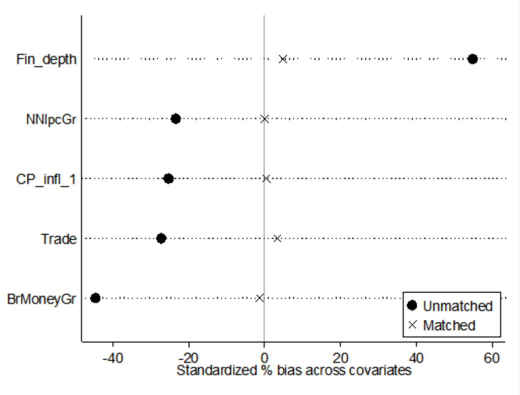


Figure 9 - Standardise Bias, Radius 0.001, Developing countries

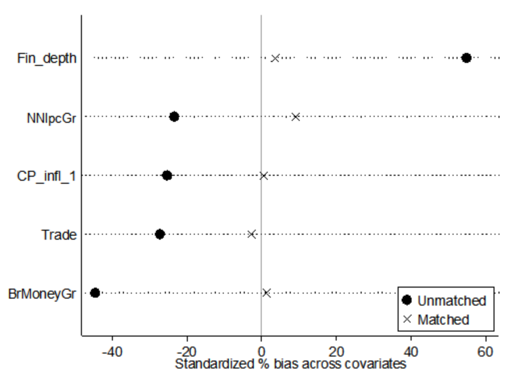


Figure 10 - Standardise Bias, Kernel Matching, Developing countries

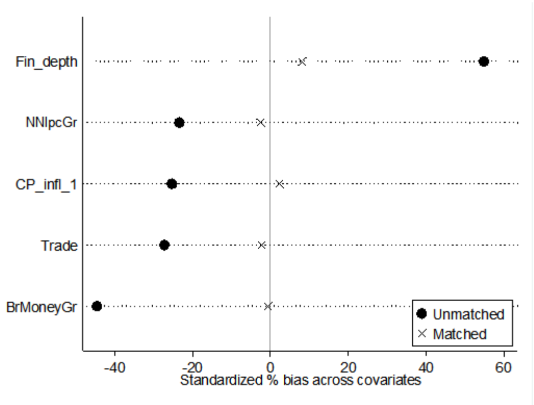


Figure 11 - Standardise Bias, Nearest Neighbor, Pooled dataset

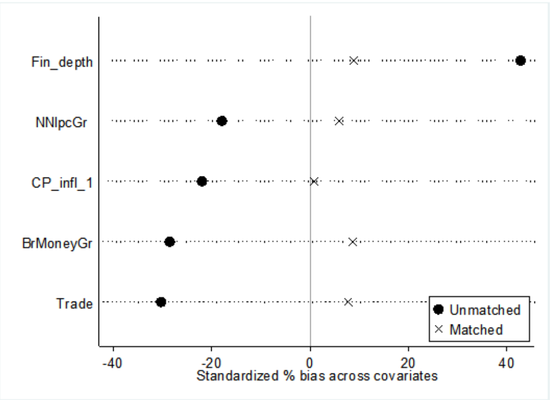


Figure 12 - Standardise Bias, 3 Nearest Neighbor, Pooled dataset

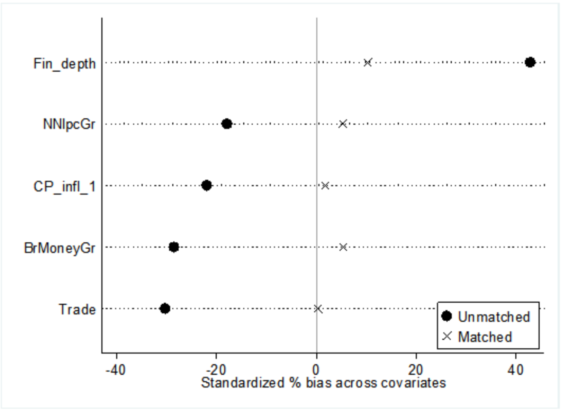


Figure 13 - Standardise Bias, Radius 0.03, Pooled dataset

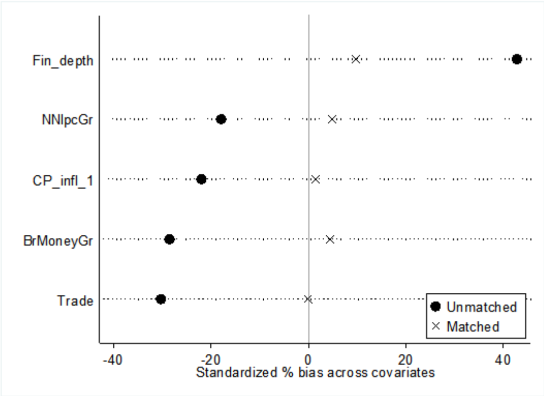


Figure 14 - Standardise Bias, Radius 0.01, Pooled dataset

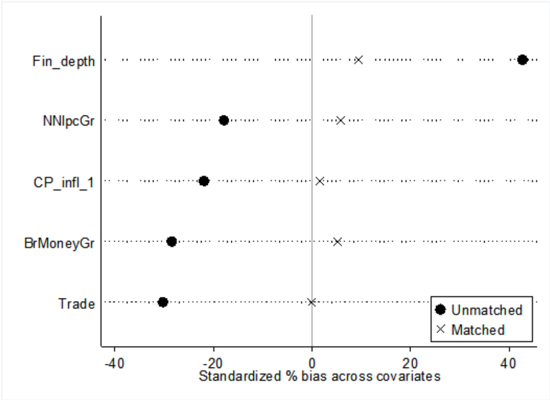


Figure 15 - Standardise Bias, Radius 0.005, Pooled dataset

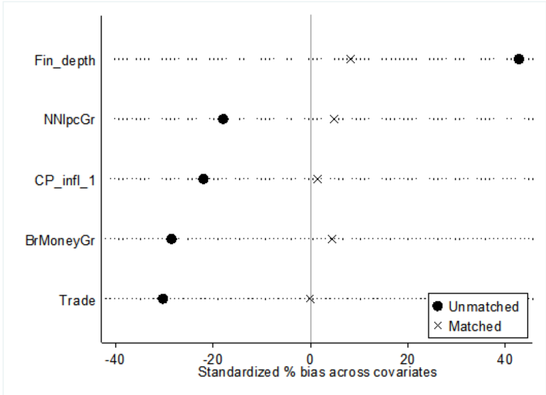


Figure 16 - Standardise Bias, Radius 0.001, Pooled dataset

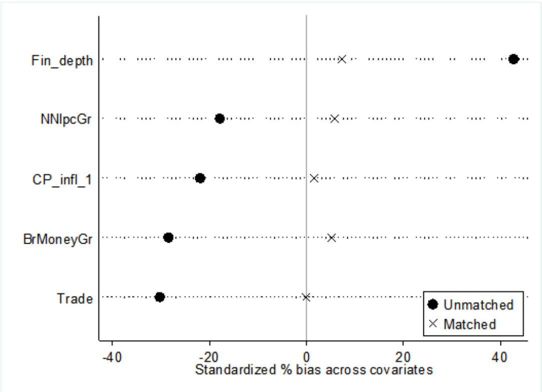


Figure 17 - Standardise Bias, Kernel Matching, Pooled dataset

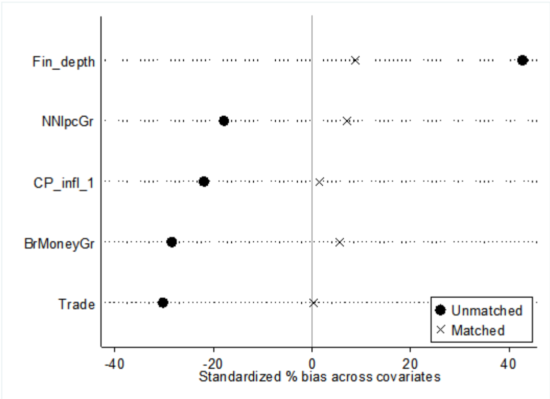


Figure 18 - Standardise Bias, Nearest Neighbor, Industrial countries

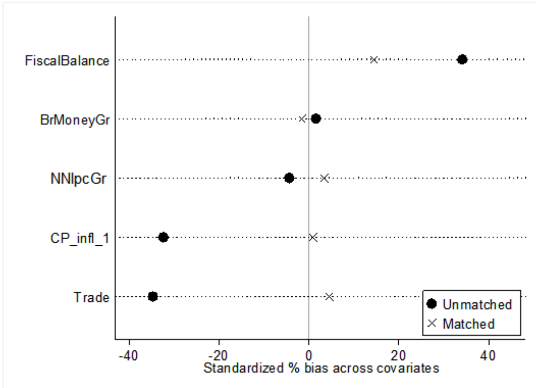


Figure 19 - Standardise Bias, 3 Nearest Neighbor, Industrial countries

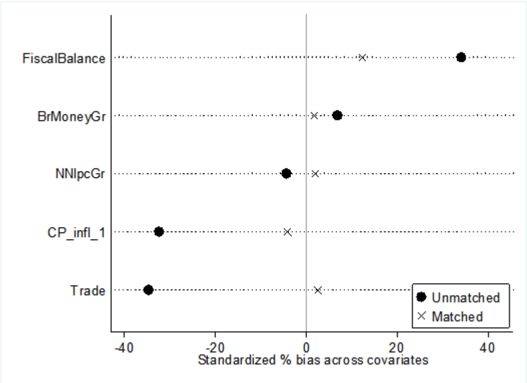


Figure 20 - Standardise Bias, Radius 0.03, Industrial countries

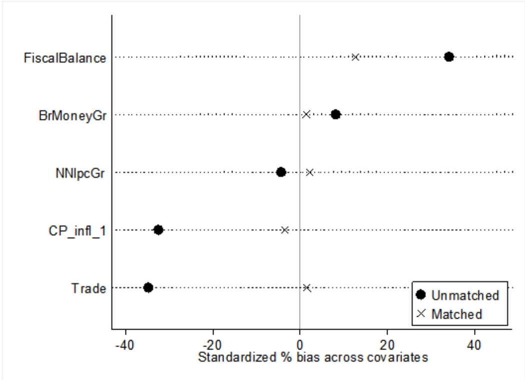


Figure 21 - Standardise Bias, Radius 0.01, Industrial countries

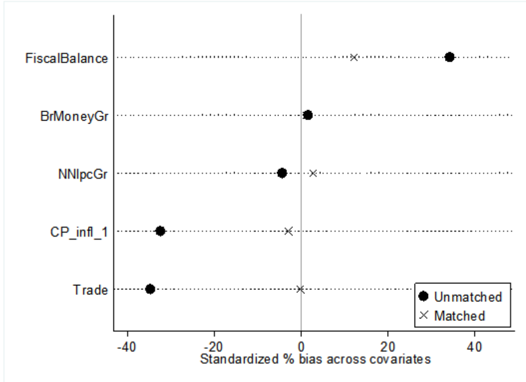


Figure 22 - Standardise Bias, Radius 0.005, Industrial countries

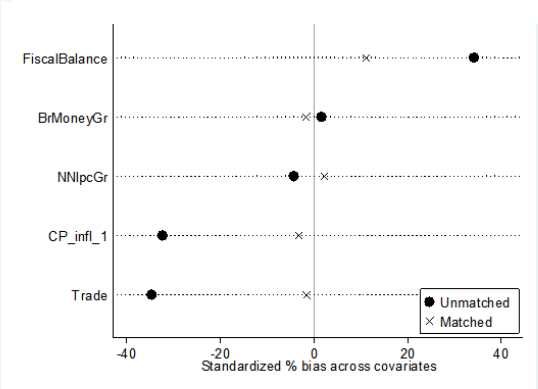


Figure 23 - Standardise Bias, Radius 0.001, Industrial countries

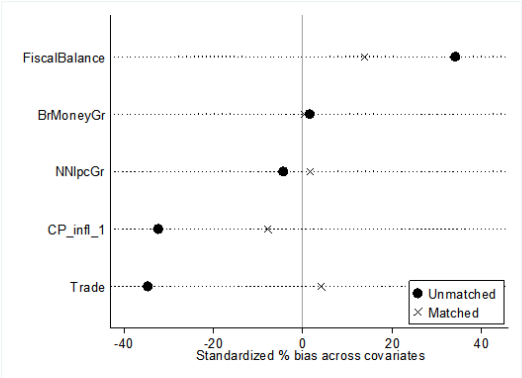


Figure 24 - Standardise Bias, Kernel Matching, Industrial countries

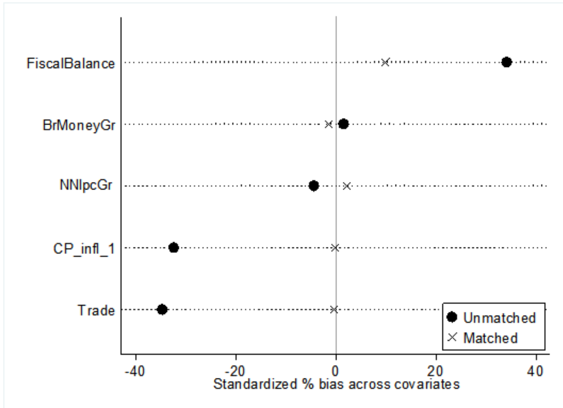


Table 29 - Robustness testing, Standardise Bias, Nearest Neighbor

			Pooled				Developing				Industrial			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	bias	Treated	Control	%bias	bias
	NNIpcGr	U	.0213	.02804	-19.6		.02425	.03349	-24.3		.01588	.01716	-5.3	
		M	.0213	.01849	8.1	58.4	.02425	.02226	5.2	78.5	.01588	.01384	8.4	-58.9
	FinD	U	.85644	.61402	51.9		.66884	.41569	73.4		1.0881	1.0103	16.6	
		M	.85644	.81353	10.1	80.5	.66884	.6374	9.1	87.6	1.0881	1.0394	10.4	37.4
Nearest	CPI_infl_1	U	.03829	.1744	-18.8		.04588	.24501	-22.6		.02606	.0333	-18.8	
Neighbor		M	.03829	.03352	0.7	96.5	.04588	.04457	0.1	99.3	.02606	.02864	-6.7	64.3
	Trade	U	.70687	.86648	-33.3		.75165	.81647	-16.2		.68091	.9664	-46.2	
		M	.70687	.72091	-2.9	91.2	.75165	.72567	6.5	59.9	.68091	.63436	7.5	83.7
	BMGr	U	.10863	.18042	-36.8		.12227	.23067	-50.7		.08156	.08001	1.3	
		M	.10863	.10029	4.3	88.4	.12227	.12127	0.5	99.1	.08156	.08232	-0.7	45.9

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.097	0.000	32.1	33.3	0.173	0.000	37.4	24.3	0.044	0.000	17.7	16.6
Matched	0.012	0.141	5.9	4.3	0.003	0.897	4.3	5.2	0.012	0.419	8.2	8.4

Table 30 - Robustness testing, Standardise Bias, 3 Nearest Neighbor

			Pooled				Developing				Industrial			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	bias	Treated	Control	%bias	bias
	NNIpcGr	U	.0213	.02804	-19.6		.02425	.03349	-24.3		.01588	.01716	-5.3	
		M	.0213	.01734	11.5	41.3	.02425	.02227	5.2	78.6	.01588	.01512	3.1	41.0
	FinD	U	.85644	.61402	51.9		.66884	.41569	73.4		1.0881	1.0103	16.6	
		M	.85644	.81051	10.3	80.1	.66884	.63622	9.5	86.7	1.0881	1.0351	10.8	34.9
3 Nearest	CPI_infl_1	U	.03829	.1744	-18.8		.04588	.24501	-22.6		.02606	.0333	-18.8	
Neighbors		M	.03829	.03444	0.5	97.2	.04588	.04344	0.3	98.8	.02606	.02983	-9.8	47.9
	Trade	U	.70687	.86648	-33.3		.75165	.81647	-16.2		.68091	.9664	-46.2	
		M	.70687	.72063	-2.9	91.4	.75165	.73041	5.3	67.2	.68091	.63932	6.7	85.4
	BMGr	U	.10863	.18042	-36.8		.12227	.23067	-50.7		.08156	.08001	1.3	
		M	.10863	.09443	7.3	80.2	.12227	.11751	2.2	95.6	.08156	.08241	-0.7	45.5

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.097	0.000	32.1	33.3	0.173	0.000	37.4	24.3	0.044	0.000	17.7	16.6
Matched	0.014	0.024	7.1	7.3	0.004	0.816	4.7	5.2	0.016	0.258	8.4	6.7

Table 31 - Robustness testing, Standardise Bias, Radius 0.03

			Pooled				**Developing**				**Industrial**			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	bias	Treated	Control	%bias	bias
	NNIpcGr	U	.0213	.02804	-19.6		.02425	.03349	-24.3		.01588	.01716	-5.3	
		M	.0213	.01904	6.5	66.6	.02425	.02479	-1.4	94.1	.01593	.014021	4.3	19.8
	FinD	U	.85644	.61402	51.9		.66884	.41569	73.4		1.0881	1.0103	16.6	
		M	.85644	.81251	10.2	80.2	.66884	.64553	6.8	90.8	1.0837	1.0254	11.2	32.4
Radius	CPI_infl_1	U	.03829	.1744	-18.8		.04588	.24501	-22.6		.02606	.0333	-18.8	
(0.03)		M	.03829	.03491	0.5	97.5	.04588	.04307	0.3	98.6	.02605	.02947	-8.9	52.7
	Trade	U	.70687	.86648	-33.3		.75165	.81647	-16.2		.68091	.9664	-46.2	
		M	.70687	.71453	-1.6	95.2	.75165	.74818	0.9	94.7	.6802	.65927	3.4	92.7
	BMGr	U	.10863	.18042	-36.8		.12227	.23067	-50.7		.08156	.08001	1.3	
		M	.10863	.10057	4.1	88.8	.12227	.12264	-0.2	99.7	.07356	.07431	-0.6	51.6

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.097	0.000	32.1	33.3	0.173	0.000	37.4	24.3	0.044	0.000	17.7	16.6
Matched	0.012	0.054	5.9	4.1	0.002	0.953	1.9	0.9	0.016	0.276	8.4	8.7

Table 32 - Robustness testing, Standardise Bias, Radius 0.01

			Pooled				**Developing**				**Industrial**			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	bias	Treated	Control	%bias	bias
	NNIpcGr	U	.0213	.02804	-19.6		.02425	.03349	-24.3		.01588	.01716	-5.3	
		M	.0213	.01852	8.0	58.9	.02455	.02268	4.9	79.8	.01593	.01506	2.3	56.1
	FinD	U	.85644	.61402	51.9		.66884	.41569	73.4		1.0881	1.0103	16.6	
		M	.85644	.80124	10.9	79.2	.64699	.61982	7.9	89.3	1.0837	1.0402	10.1	34.5
Radius	CPI_infl_1	U	.03829	.1744	-18.8		.04588	.24501	-22.6		.02606	.0333	-18.8	
(0.01)		M	.03829	.03417	0.6	97.0	.04596	.04218	0.4	98.1	.02605	.02936	-8.6	54.3
	Trade	U	.70687	.86648	-33.3		.75165	.81647	-16.2		.68091	.9664	-46.2	
		M	.70687	.71328	-1.3	96.0	.75128	.75561	-1.1	93.3	.6802	.64523	5.7	87.7
	BMGr	U	.10863	.18042	-36.8		.12227	.23067	-50.7		.08156	.08001	1.3	
		M	.10863	.09941	4.7	87.2	.12427	.11984	2.1	95.9	.07356	.07408	-0.4	66.7

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.097	0.000	32.1	33.3	0.173	0.000	37.4	24.3	0.044	0.000	17.7	16.6
Matched	0.012	0.040	6.0	4.7	0.004	0.844	3.3	2.1	0.017	0.231	7.9	7.8

Table 33 - Robustness testing, Standardise Bias, Radius 0.005

			Pooled				Developing				Industrial			
		Unmatched	Mean		%bias	%reduct	Mean		%bias	%reduct	Mean		%bias	%reduct
Matching	Variable	Matched	Treated	Control		bias	Treated	Control		bias	Treated	Control		bias
	NNIpcGr	U	.0213	.02804	-19.6		.02425	.03349	-24.3		.01588	.01716	-5.3	
		M	.0213	.01894	6.8	65.1	.02458	.02127	8.7	64.2	.01593	.01423	7.0	-32.4
	FinD	U	.85644	.61402	51.9		.66884	.41569	73.4		1.0881	1.0103	16.6	
		M	.85644	.79514	11.0	79.2	.6303	.60727	6.7	90.9	1.0837	1.0297	11.1	33,2
Radius	CPI_infl_1	U	.03829	.1744	-18.8		.04588	.24501	-22.6		.02606	.0333	-18.8	
(0.005)		M	.03829	.03403	0.6	96.9	.04633	.04244	0.4	98.0	.02605	.02884	-7.3	61.5
	Trade	U	.70687	.86648	-33.3		.75165	.81647	-16.2		.68091	.9664	-46.2	
		M	.70687	.71312	-1.3	96.1	.74941	.767	-4.4	72.9	.6802	.64811	5.2	88.8
	BMGr	U	.10863	.18042	-36.8		.12227	.23067	-50.7		.08156	.08001	1.3	
		M	.10863	.10035	4.2	88.5	.12505	.11841	3.1	93.9	.07356	.07409	-0.4	65.7

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.097	0.000	32.1	33.3	0.173	0.000	37.4	24.3	0.044	0.000	17.7	16.6
Matched	0.012	0.047	5.6	4.2	0.005	0.783	4.7	4.4	0.015	0.297	8.1	7.0

Table 34 - Robustness testing, Standardise Bias, Radius 0.001

			Pooled				Developing				Industrial			
		Unmatched	Mean		%bias	%reduct	Mean		%bias	%reduct	Mean		%bias	%reduct
Matching	Variable	Matched	Treated	Control		bias	Treated	Control		bias	Treated	Control		bias
	NNIpcGr	U	.0213	.02804	-19.6		.02425	.03349	-24.3		.01588	.01716	-5.3	
		M	.02156	.01868	8.4	57.2	.02595	.02391	5.4	77.8	.01619	.01503	4.8	9.3
	FinD	U	.85644	.61402	51.9		.66884	.41569	73.4		1.0881	1.0103	16.6	
		M	.83905	.80146	10.5	79.5	.51579	.49434	6.2	91.5	1.0863	1.0302	10.9	33,6
Radius	CPI_infl_1	U	.03829	.1744	-18.8		.04588	.24501	-22.6		.02606	.0333	-18.8	
(0.001)		M	.03844	.03446	0.6	97.1	.04919	.04779	0.2	99.3	.02555	.02829	-7.1	62.3
	Trade	U	.70687	.86648	-33.3		.75165	.81647	-16.2		.68091	.9664	-46.2	
		M	.70802	.71581	-1.6	95.1	.75593	.74649	2.4	85.4	.69413	.66797	4.2	90.8
	BMGr	U	.10863	.18042	-36.8		.12227	.23067	-50.7		.08156	.08001	1.3	
		M	.10758	.09642	5.7	84.5	.12754	.12483	1.3	97.5	.07017	.06954	0.4	67.4

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.097	0.000	32.1	33.3	0.173	0.000	37.4	24.3	0.044	0.000	17.7	16.6
Matched	0.013	0.042	6.1	5.7	0.002	0.973	3.1	2.4	0.016	0.322	9.0	6.9

Table 35 - Robustness testing, Standardise Bias, Kernel matching

			Pooled				Developing				Industrial			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	bias	Treated	Control	%bias	bias
	NNIpcGr	U	.0213	.02804	-19.6		.02425	.03349	-24.3		.01588	.01716	-5.3	
		M	.0213	.01663	13.5	30.8	.02425	.02098	8.8	63.7	.01588	.01481	4.9	7.5
	FinD	U	.85644	.61402	51.9		.66884	.41569	73.4		1.0881	1.0103	16.6	
		M	.85644	.80201	10.4	79.8	.66884	.61654	7.1	79.3	1.0881	1.0125	16.2	2.8
Kernel Matching	CPI_infl_1	U	.03829	.1744	-18.8		.04588	.24501	-22.6		.02606	.0333	-18.8	
		M	.03829	.03366	0.6	96.6	.04588	.04387	0.2	99.0	.02606	.02705	-2.6	86.2
	Trade	U	.70687	.86648	-33.3		.75165	.81647	-16.2		.68091	.9664	-46.2	
		M	.70687	.72505	-3.8	88.6	.75165	.73021	6.4	66.1	.68091	.65553	4.1	91.1
	BMGr	U	.10863	.18042	-36.8		.12227	.23067	-50.7		.08156	.08001	1.3	
		M	.10863	.09706	5.9	83.9	.12227	.11924	1.4	97.2	.08156	.08196	-0.3	90.3

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.097	0.000	32.1	33.3	0.173	0.000	37.4	24.3	0.044	0.000	17.7	16.6
Matched	0.016	0.011	7.7	5.9	0.008	0.449	4.8	10.4	0.013	0.400	8.7	10.2

Table 36 - Standardise Bias, Inverse Probability Weighting

			Pooled				Developing				Industrial			
		Unmatched	Mean			%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control	%bias	bias	Treated	Control	%bias	bias	Treated	Control	%bias	bias
	NNIpcGr	U	.02237	.02868	-17.9		.01103	.01469	-21.8		.0164	.01697	-2.3	
		M	.02237	.01946	8.2	53.9	.01103	.01015	5.3	75.9	.0164	.01677	-1.5	34.3
	FinD	U	.82985	.62644	42.8		.61291	.42587	54.8		1.1973	.97405	48.3	
		M	.82985	.73821	19.3	54.9	.61291	.58632	7.8	85.8	1.1973	1.2018	-1.0	98.0
Inverse Probability Weighting	CPI_infl_1	U	.04005	.09321	-22.0		.04946	.12529	-25.3		.02399	.03761	-33.1	
		M	.04005	.03794	0.9	96.0	.04946	.04618	1.1	95.7	.02399	.02485	-2.1	93.7
	Trade	U	.72132	.86086	-30.2		.74245	.84906	-27.2		.68676	.8813	-34.6	
		M	.72132	.69791	5.1	83.2	.74245	.73678	1.4	94.7	.68676	.66911	3.1	90.9
	BMGr	U	.1107	.15041	-28.5		.1275	.19067	-44.6		.08292	.08064	1.9	
		M	.1107	.10045	7.4	74.2	.1275	.12967	-1.5	96.6	.08292	.08175	0.9	52.1

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.072	0.000	28.3	28.5	0.135	0.000	34.8	27.2	0.065	0.000	24.0	33.1
Matched	0.015	0.005	8.2	7.4	0.004	0.770	3.4	1.5	0.001	0.995	2.1	2.1

Table 37 - Robustness testing, Standardise Bias, Inverse Probability Weighting

			Pooled				Developing				Industrial			
		Unmatched	Mean		%bias	%reduct	Mean			%reduct	Mean			%reduct
Matching	Variable	Matched	Treated	Control		bias	Treated	Control	%bias	bias	Treated	Control	%bias	bias
Inverse Probability Weighting	NNIpcGr	U	.00902	.01165	-17.9		.01025	.01386	-22.2		.00674	.00724	-4.8	
		M	.00902	.00684	14.8	17.4	.01025	.00811	13.2	40.8	.00674	.0057	10.0	-107.2
	FinD	U	.85644	.61402	51.9		.66884	.41569	73.4		1.0881	1.0103	16.6	
		M	.85644	.78797	14.7	71.8	.66884	.61654	15.2	79.3	1.0881	1.0125	16.2	2.8
	CPI_infl_1	U	.03829	.1744	-18.8		.04588	.24501	-22.6		.02606	.0333	-18.8	
		M	.03829	.03366	0.6	96.6	.04588	.04387	0.2	99.0	.02606	.02705	-2.6	86.2
	Trade	U	.70687	.86648	-33.3		.75165	.81647	-16.2		.68091	.9664	-46.2	
		M	.70687	.72505	-3.8	88.6	.75165	.71021	10.4	36.1	.68091	.65553	4.1	91.1
	BMGr	U	.10863	.18042	-36.8		.12227	.23067	-50.7		.08156	.08001	1.3	
		M	.10863	.09706	5.9	83.9	.12227	.11924	1.4	97.2	.08156	.08206	-1.4	-9.1

Sample	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias	Ps R2	p>chi2	MeanBias	MedBias
Unmatched	0.097	0.000	31.7	33.3	0.172	0.000	37.0	22.6	0.044	0.000	17.6	16.6
Matched	0.017	0.008	8.0	5.9	0.009	0.388	8.1	10.4	0.012	0.406	6.7	10.0