What Happens When Performance Based Financing Meets Free Healthcare? Evidence From An Interrupted Time-Series Analysis With Independent Controls

Corresponding author's contact information:
Naasegnibe Kuunibe (corresponding author)
Affiliation: Heidelberg Institute for Global Health, University Hospital and Medical Faculty, Heidelberg University, Heidelberg, Germany
Address: Im Neuenheimer Feld 365, 69120 Heidelberg, Germany
Telephone number: +49 15219302268
Email: nkuunibe@uds.edu.gh or naasegnibe.kuunibe@uni-heidelberg.de

Author's affiliation and qualifications:

1. Naasegnibe Kuunibe (corresponding author), PhD candidate
   Affiliation: Heidelberg Institute of Global Health, Medical Faculty, Heidelberg University, Heidelberg, Germany and University for Development Studies, Tamale, Ghana
   Email: nkuunibe@uds.edu.gh or naasegnibe.kuunibe@uni-heidelberg.de
   ORCID: 0000-0003-1975-3718

2. Julia Lohmann, PhD
   Affiliation: Heidelberg Institute of Global Health, University Hospital and Medical Faculty, Heidelberg University, Heidelberg, Germany and London School of Hygiene and Tropical Medicine, UK
   Email: julia.lohmann@uni-heidelberg.de
   ORCID: 0000-0003-4136-9296

3. Michael Hillebrecht, PhD
   Affiliation: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ)
   Email: Michael.hillebrecht@giz.de

4. Hoa Thi Nguyen, PhD candidate
   Affiliation: Heidelberg Institute of Global Health, University Hospital and Medical Faculty, Heidelberg University, Heidelberg, Germany
   Email: hoa.nguyen@uni-heidelberg.de

5. Gauthier Tougri, MD
   Affiliation: Burkina Faso Ministry of Health
   Email: tougriprimus@yahoo.fr

6. Manuela De Allegri, MSc, PhD
   Affiliation: Heidelberg Institute of Global Health, University Hospital and Medical Faculty, Heidelberg University, Heidelberg, Germany
   Email: manuela.deallegri@uni-heidelberg.de
   ORCID : 0000-0002-8677-1337

Keywords:
Performance-based financing, control interrupted time series analysis, multiple imputation, Burkina Faso

JEL classification: 118, C22

Acknowledgements: We are grateful to the Ministry of Health of Burkina Faso for granting us access to the data, to members of the PBF Technical Service within the Ministry of Health for providing insight into the intervention design and implementation process, and to Paul Jacob Robyn and Christophe Rockmore for their valuable comments on an earlier draft of this manuscript.

Funding:The analysis presented in this manuscript was conducted by the authors outside the scope of the World Bank HRITF-funded impact evaluation. All authors contributed their own time towards preparation of this article without direct funding from the HRITF for preparation of this article. Naasegnibe Kuunibe received a PhD fellowship from the Ghana Ministry of Education in collaboration with the German Academic Exchange Service (DAAD).

Competing interests

The analysis presented in this manuscript was conducted outside the framework of the primary Impact Evaluation funded by the Health Results Innovation Trust Fund (HRITF) of the World Bank, with all authors contributing their own time and no funding agency playing any role in the design of the study, the analysis of the data, or the writing of the manuscript.

Nevertheless, for full transparency, we wish to note that Manuela De Allegri served as the Principal Investigator of the primary Performance Based Financing Impact Evaluation conducted in Burkina while Julia Lohmann and Michael Hillebrecht were part of the Impact Evaluation team. None of the authors received any direct payment from the World Bank, since the research grant was managed by the administration of the University Hospital and Medical Faculty.

Authors' contributions

Naasegnibe Kuunibe, Julia Lohmann and Manuela De Allegri conceived the study. Nassegnibe Kuunibe, Julia Lohmann , Gauthier Tougri and Michael Hillebrecht contributed to acquisition and management of data. Naasegnibe Kuunibe undertook data analysis with support from Julia Lohmann, Hoa Thi Nguyen and Manuela De Allegri. All authors contributed to the interpretation of the results. Naasegnibe Kuunibe drafted the manuscript with contributions from all authors. All authors read and approved the final manuscript.

Ethical consideration

This study presents no major ethical challenges since it is based on an analysis of fully anonymous secondary data. The study received ethical approval from the Ethics Commission of the Medical Faculty of the University of Heidelberg (approval number S-272/2013) and the Ethics Committee of the Burkina Faso Ministry of Health (approval number 2017-9-138). In addition, the study obtained written permission from the Burkina Faso Ministry of Health to download and use the data for the analysis.

Abstract

Introduction: In spite of the wide attention Performance-based financing (PBF) has received over the past decade, no evidence is available on its impacts on quantity and mix of service provision nor on its interaction with parallel health financing interventions. Our study aimed to examine the PBF impact on quantity and mix of service provision in Burkina Faso, while accounting for the parallel introduction of a free healthcare policy.

Methods: We used Health Management Information System (HMIS) data from 838 primary-level health facilities across 24 districts and relied on an interrupted time series analysis with independent controls. We placed two interruptions, one to account for PBF and one to account for the free healthcare policy.

Results: In the period before the free healthcare policy, PBF produced significant but modest increases across a wide range of maternal and child services, but a significant decrease on child immunization coverage. In the period after the introduction of the free healthcare policy, PBF did not affect service provision in intervention compared to control facilities, possibly indicating a saturation effect.

Conclusion: Our findings indicate that PBF can produce modest increases in service provision, without altering the overall service mix. Our findings, however, also indicate that the introduction of other health financing reforms can quickly crowd out the effects produced by PBF. Further qualitative research is required to understand what factors allow healthcare providers to increase the provision of some, but not all services and how they react to the joint implementation of PBF and free healthcare.

1. Introduction

Over the last decade, performance-based financing (PBF) has been introduced in over 30 low- and middle-income countries (LMICs) to strengthen health service provision [1-3]. PBF is a supply-side intervention aiming to motivate healthcare providers to improve their performance by providing financial incentives for the delivery of predefined services while concurrently enhancing supportive supervision and autonomy in financial management [1, 4]. In Sub-Saharan Africa (SSA), a number of robust primary data-based impact evaluations (IEs) have been conducted to estimate PBF effects on health service utilization [5-7] and, to a lesser extent, on the quality of service provision [8, 9]. Emerging evidence, however, continues to be mixed regarding both the direction and the size of effects [5, 8, 10, 11].

There are many advantages to primary data-based IEs [12], but they are also costly and usually based only on a limited number of data points. Follow-up is usually conducted shortly after the introduction of an intervention (for pragmatic reasons), so longer-term effects are often not evaluated [13, 14]. To address these weaknesses, researchers have recently proposed [14] to turn to routine secondary data to complement primary data-based IEs. Impact analyses based on secondary data allow researchers to evaluate the immediate and long-term impacts of interventions. Recent applications include two studies in Burkina Faso that evaluated the impact of user fee exemptions for women at birth [15] and for children under five [16]. Within the field of PBF, two studies in Burkina Faso employed secondary data to evaluate PBF impacts. The first study used Health Management Information System (HMIS) data and employed a difference-in-differences analytical approach to look at effects of a small-scale PBF pilot implemented in three districts on the utilisation of maternal health services [17]. Also relying on HMIS district-level indicators, the second study was conducted half-way through an extension of the PBF pilot and examined the program's contribution to improving maternal and child healthcare indicators [18]. Studies from other countries in Sub-Saharan Africa include two in Rwanda; one which employed quarterly routine facility data from 2006 to 2010 to describe PBF contribution to the quality of service delivery over a five-year period [19]; and one relying on monthly routine facility data to assess the impact of PBF on quality of care over a three-year period [20]. In Burundi, a study employed HMIS data aggregated at the province level to evaluate the effect of PBF on use of primary healthcare services in the context of selective free healthcare provision [21].

In Burkina Faso, PBF was first piloted in 2011 in three out of 63 health districts to address the inadequate provision of maternal and child health (MCH) services. In 2014, the government extended the program to another 12 districts, covering MCH and general curative services [22]. Approximately two years into the scale-up, in June 2016, the government introduced a nationwide free healthcare policy for MCH services (hereafter referred to as *gratuité*) [23, 24], implemented in parallel with PBF.

A primary data-based IE, funded by the World Bank's Health Results Innovation Trust Fund (HRITF) and spanning the period from 2014 to early 2017, showed a positive impact of PBF on the utilisation of facility-based delivery and post-natal care, but not on other key MCH service utilisation indicators [25]. The evaluation relied on a pre- and post-test design with independent controls using a difference-in-differences design. With data being collected only at two-time points (2013 and 2017), it was impossible for the impact evaluation to ascertain any specific interaction between PBF and *gratuité*. Similarly, the above-mentioned recent evaluation based on district-level HMIS data for the period 2013 to 2016 detected a positive impact of PBF only on postnatal care but did not take into account the fact that *gratuité* was introduced mid-way through the PBF implementation [18].

Using facility-level HMIS data from 24 districts for the period January 2013 to September 2017, our study aims at complementing the evidence emerging from the above-mentioned studies by assessing the impact of PBF on nine MCH and curative care indicators. We employed a quasi-experimental design, interrupted time series analysis (ITSA) with independent controls, to overcome the drawbacks inherent in primary data-based evaluations and to account explicitly for the introduction of the *gratuité* mid-way through the PBF implementation. It ought to be noted that while our models include the *gratuité*, they do so exclusively to provide a more accurate estimate of PBF impact and not to estimate the impact of *gratuité per se*. Moreover, our work does not focus on the mechanisms of interaction between the two programs.

2. Methods

2.1. Study settings

Burkina Faso is a West-African landlocked country with an estimated population of about 18 million people (2015) [26], about half of which live in poverty [27]. On the 2017 Human

Development Index (HDI) Burkina Faso ranked 183 out of 189. The country's gross national income (GNI) per capita was 1,650 USD (PPP) in 2017 [28].

In 2014, total expenditure on health as a percentage of GDP was only 5%. Government health expenditure as a percentage of total health expenditure was only 11.2%, while private expenditure (mostly through user fees at point of service) constituted 47.7% of total health expenditure [27]. The lack of adequate funding for health continues to pose a major challenge to any effort made towards Universal Health Coverage [29]. Over the last few years, the country saw some improvement in MCH indicators with under-five mortality and maternal mortality dropping respectively from 186 per 1,000 live births in 2000 to 114 per 1,000 live births in 2010 [30] and from 580 per 100,000 live births in 2000 to 330 per 100,000 live births in 2015 [28], but mortality levels remain high in absolute terms.

Health service delivery is organised mostly through the public sector along a three-tier system. The first level consists of the "*Centres de Santé et de Promotion Sociale"* (CSPS) and district hospitals called *"Centres médicaux avec antenne chirurgicale"* (CMA), which provide basic preventive and curative care services to the community. The second level consists of Regional Hospitals, which serve as referral centres for CMAs, while specialist third-level care is provided in University Hospital Centres [31].

2.2. PBF program in Burkina Faso

Against this background, the Ministry of Health of Burkina Faso, with support from the HRITF, implemented an extended PBF pilot in 12 out of its 60 districts with the specific objective of improving MCH indicators. The 12 intervention districts were purposely selected by the government and its development partners for its particularly poor health indicators [32]. During this extended PBF pilot, healthcare facilities in the 12 intervention districts were rewarded based on verified quantity and quality indicators of service provision. Rewards attached to quantity were computed based on pre-defined unit prices, while rewards attached to quality were computed as a function of the quantity reward, provided the facility could prove that it had attained pre-defined quality targets (See Appendix A for a list of indicators). The accompanying IE introduced above relied on a pre- and post-test design with independent controls, comparing health service utilization and quality indicators across the 12 intervention districts and 12 control districts. Control districts were also selected in consultation between government, development partners, and researchers to be as comparable as possible, in terms of overall socio-economic status and health system structures, to the intervention ones. In selected facilities in 10 out of these 12 intervention districts, PBF was accompanied by a number of equity measures aimed at testing the combined effect of the two interventions in reaching out to specifically vulnerable populations. Futher details are available elsewhere [32].

In this study, we do not differentiate facilities according to whether they did or did not implement an equity measure as we focus on the supply-side intervention to establish whether PBF produced changes in quantity and mix of service provision, irrespective of the distributional incidence of these changes which is addressed by a parallel study [33].

2.3. Study design

This study employed an interrupted time series analysis (ITSA) design with independent controls [34, 35]. Although we conducted the analysis independently of the primary IE, we made use of the IE design, using data from the same 12 intervention and 12 control districts.

Recognized as the best approach to impact assessment when randomisation is not possible [36, 37], the ITSA design allowed us to discern the effects of secular time trends from that of PBF by accounting for existing trends before its introduction [38, 39]. The control group further allowed to control for non-program influences occurring in concurrence to PBF to which both intervention and control facilities were exposed, such as *gratuité* [34]. Thus, we evaluated PBF impacts by looking at whether the intervention facilities deviated from baseline levels and trends to a greater extent than the control facilities [40].

2.4. Conceptual framework

The inclusion of *gratuité* in the model was guided by a theoretical understanding that the two interventions may interact, and that the impact of PBF could therefore only be estimated by taking into account the introduction of the *gratuité*. We start our theoretical reflection by looking at the expected theory of change in each of the two interventions separately.

On the one hand, in line with principal-agent theory [41, 42], we postulated that the introduction of financial incentives, the increased supportive supervision and the enhanced managerial autonomy promoted by PBF (without *gratuité*) would motivate health workers to align their behaviour with Ministry of Health expectations, resulting in increased provision of quality healthcare services. On the other hand, in line with the literature emerging from pilot experiences [15-18], we postulated that the removal of user payments at the point of use for MCH services promoted by the *gratuité* would lift financial barriers to access and stimulate demand. We further postulated that in the period before *gratuité*, all other financial arrangements held equal and hence reacting exclusively to PBF (including financial incentives, increased supervision, and enhanced autonomy), health workers might have privileged the provision of some services above others. For instance, we expected providers to favor services generating greater financial gains and/or services whose provision could more easily be directed by their own behavior independently of demand. In turn, we expected this preference for selected services to lead to important alterations in the quantity and the mix of service provision, fully driven by the supply side. We expected that the introduction of the *gratuité* alongside PBF (PBF+ *gratuité*) would alter PBF dynamics by removing user charges only for a subset of services included in the PBF package (i.e. MCH services).

In fact, following the introduction of the *gratuité*, providers received dual reimbursement for the services (maternal care and child curative services) included both in the PBF and in the *gratuité* package [43, 44]. Hence, we expected the *gratuité* to increase the provision of this specific sub-set of services through two parallel mechanisms. On the one side, we expected targeted exemptions to stimulate demand for this specific sub-set of services, potentially altering the quantity and composition of the patient population. In turn, we expected this shift to enable health providers to increase service provision also for services more closely tied to demand-side barriers, where the change in supply-side processes promoted by PBF might have not been sufficient to generate change. On the other side, we expected that the dual reimbursement would encourage providers to keep up and further encourage service provision of the selected sub-set of services included in both the PBF and the *gratuité* package. Hence, by acting to modify both demand- and supply-side, the interaction between PBF and *gratuité* should lead to important shifts in service provision quantity and mix beyond what PBF could effect alone.

2.5. Data sources and variables

2.5.1. Data and sources

The study used data from the HMIS and the facility catchment area population. In line with the PBF intervention focus, we focused the analysis exclusively on public primary level facilities and worked with data from 838 (524 intervention and 314 control) CSPS for which we had sufficiently complete data from both datasets and for all relevant indicators for the period from January 2013 to September 2017, representing 12 months pre-intervention, 29 months post-PBF introduction and 16 months post-PBF+*gratuité*. HMIS data were organised to report monthly patient counts on each services, while catchment area population data comprised annual projections of the total population within the facility catchment area, the population aged less than one year, the population aged less than five years, women of reproductive age, expected pregnancies and expected deliveries. While no specific independent quality assessments are available, HMIS data in Burkina Faso are generally considered of good quality and have been used extensively for scientific purposes over the past few years [15, 16, 18]).

It is worthwhile noting that although we excluded *a priori 221* facilities with missing data on more than 80 percent of HMIS data points (ie facility-months per indicator), our sample still included a number of missing values, with proportions of missing values ranged from less than 7% for OPD to 33.5% for short-term family planning service. Most of the missings were due to facilities having been newly created in the course of the time period under consideration in this paper. Most missings are therefore structural rather than an indication of poor data quality. In line with existing literature [45, 46], we adopted multiple imputations in response. First, we did mean imputation for delivery services (a relatively stable indicator in Burkina Faso[1]) and used the mean imputed version as our independent variable (since all indicators had missing values and were not suitable to be used as independent variables) to impute the other indicators. Five rounds of imputations were done in Stata using the mi impute Poisson command [47] and the resulting average values were then used for analysis.

2.5.2. Outcome variables

We focused our analysis on the ten (out of a total of 23 as in Appendix A) PBF indicators for which sufficient numbers of patients were recorded in the HMIS database. This exclusion is in no way related to actual program impact on service provision but reflects the fact that some conditions and hence some treatments (e.g. tuberculosis, HIV), although incentivized in the context of PBF, occur too rarely in Burkina Faso to generate sufficient data for analysis. The ten indicators are referred to as outcome variables in the following. To produce values that can be interpreted from a policy perspective, we constructed each outcome variable to represent the ratio of services provided by a CSPS to people living within its catchment area in a given month (and interpreted as percentages). Table 1 reports the numerator and denominator for each outcome variable

[1] This was the outcome variable that did not show AR autocorrelation at lag 1. This may be due to the consistent policy intervention on the indicator in the past

included in the study. For analysis, we aggregated data into monthly averages, differentiating control from intervention facilities.

2.5.3. Time and program variables

We counted time in months starting January 2013 and coded serially from one to 57, representing September 2017. Given the aim of this paper was to assess the impact of PBF only, accounting for the implementation of gratuite, we included set of dummy variables to capture the period when PBF only (without gratuite) was implemented and then took into account when the gratuite was implemented alongside PBF. Thus, we included two dummy variables to capture time in relation to PBF and the *gratuité*. The first variable identified the pre-intervention months (0=no PBF and no *gratuité* in place (January 2013 to December 2013); 1=PBF only in place (January 2014 to May 2016)); the second variable identified months during which both PBF and *gratuité* were in place (0=pre-PBF+*gratuité* in place (January 2013 to May 2016); 1=both PBF+*gratuité* in place (June 2016 to September 2017). We also included a third dummy variable that identified PBF assignment (0=control facilities; 1=intervention facilities). In addition, we included interaction terms to capture the PBF-only pre-intervention trend in control facilities, the difference in PBF-only pre-intervention trend between intervention and control facilities, the difference in level between intervention and control facilities in the period immediately following PBF-only, and the difference in PBF-only post-intervention trend between intervention and control facilities. We also included interaction terms that captured PBF+*gratuité* pre-intervention level in the control facilities, PBF+*gratuité* pre-intervention trend in control facilities, the difference in the level immediately following PBF+*gratuité*, and the difference in PBF+*gratuité* post-intervention.

2.6. Empirical analysis

First, we used scatter plots to inspect data for each outcome variable and identify outliers. Then we inspected the data for autocorrelation by using the autocorrelation (ACF) and partial autocorrelation (PACF) plots. We also examined the data for trend stationarity using the Dicky-Fuller (DF) test by decomposing the trend into systematic, periodic, and random components. We used the Akaike information criterion (AIC) and Bayesian information criterion (BIC) to select the appropriate functional forms (between linear, log-linear and the quadratic functional forms) [48] before estimating the regression equations.

Preliminary checks revealed the presence of autocorrelation for all outcome variables except delivery care. In addition, the test for model functional form (see Table 2), using the minimum values of both the AIC and the BIC, showed that the quadratic functional form best fit all indicators except delivery care, family planning, and growth monitoring for children aged between 12 to 23 months. For these three variables, AIC and BIC values detected only a small difference between the linear and the quadratic functional forms. Consequently, we implemented the quadratic functional form for each outcome variable in the regression models explained in the following [48]**.**

Second, we employed the segmented regression model [13, 49, 50] in equation (1),

$$Y_t = \theta_0 + \theta_1 T_t + \theta_2 X_t + \theta_3 X_t T_t + \theta_4 Z + \theta_5 Z T_t + \theta_6 Z X_t + \theta_7 Z X_t T_t + \theta_8 G + \theta_9 G T_t + \theta_{10} G X_t + \theta_{11} G X_t T_t + \emptyset T^2 + \mu_t \quad (1)$$

where Y_t is the rate (proportion) of services provided to people living in a facility catchment area, for each of our ten outcome variables. The following coefficients denote the PBF-only effect as follows: θ_0 is the pre-intervention level in control facilities, θ_1 is the pre-intervention trend in control facilities, θ_2 is the immediate impact in control facilities and θ_3 is the post-intervention trend in control facilities. θ_4 is the difference in pre-intervention level, θ_5 is the difference in pre-intervention trend, θ_6 is the difference in level in the period immediately following PBF and θ_7 is the difference in PBF post-intervention trend. The remaining coefficients denote the PBF+*gratuité* effect as follows: θ_8 is PBF+*gratuité* pre-intervention level in control facilities, θ_9 is PBF+*gratuité* pre-intervention trend in control facilities, θ_{10} is the difference in level immediately following *gratuité*, θ_{11} is the difference in trend in the PBF+*gratuité* post-implementation period. The coefficient $\emptyset$ is the effect of the quadratic term. We used the imputed data for the main analysis and the original data for sensitivity analyses. We estimated equation (1) separately for each outcome variable, for a total of twenty equations - ten for the main analysis and ten for the sensitivity analysis. Results of the sensitivity analyses, presented in Appendices B and C, did show any PBF effect for most indicators. This was most probably due to the use of data with some missing values [46, 51], and justifies our decision to do multiple imputation before using data for our main analysis.

Since autocorrelation was detected, we used the generalised linear model (GLM) to adjust for the presence of autocorrelation, and calculated the unadjusted and adjusted Durbin-Watson (DW) statistics to ensure correction for autocorrelation.

3. Results

3.1. Average provision of healthcare services

Table 3 presents average service provision for intervention and control facilities for the pre-intervention period (January 2013 to December 2013); the period when PBF-only was in place (January 2014 to May 2016); the period when *gratuité* was implemented alongside PBF (June 2016 to September 2017); and the entire study period (January 2013 to September 2017). Results indicate that in the 12 months preceding PBF implementation in intervention facilities, the average monthly rate of service provision per facility catchment area varied from a low of 2.5% to a high of 24% for family planning and ANC respectively. In control facilities, the average service provision ranged from a low of 2.5% for family planning services to a high of 25% for ANC. During the 29 months (January 2014 to May 2016) when PBF alone was implemented, the average monthly rate of service provision in intervention facilities varied from a low of 2.6% to a high of 24% for family planning services and ANC respectively. In control facilities, service provision varied from a low of 2.7% for family planning services to a high of 24% for ANC. When *gratuité* was implemented alongside PBF (from June 2016 to September 2017) the average monthly rate of service provision increased for most indicators in both intervention and control facilities, when compared with the period where PBF alone was implemented. For the entire study period, average service provision for most indicators was higher in intervention facilities than in the pre-intervention period when PBF alone was implemented but lower than when *gratuité* was implemented alongside PBF. In control facilities, the average for most indicators for the entire study period was lower than in the pre-intervention period, the period when PBF alone was implemented as well as in the period when *gratuité* was implemented alongside PBF.

3.2. Pre-intervention comparison

Tables 4 and 5 present results from the ITSA analysis. Corresponding graphs are in Figures 1 and 2. The unadjusted DW statistics indicate the presence of autocorrelation, while the adjusted DW statistics indicate that adjusting for autocorrelation at lag 1 was successful in addressing the autocorrelation issue. Results from the models indicate that at baseline (the period before January 2014), there was no significant difference between control and intervention facilities in the levels of service provision for all outcome variables except for tetanus vaccination in pregnancy for which we detected a significant negative difference. This means that before PBF, control facilities provided tetanus doses to more pregnant women than intervention facilities.

Similarly, before January 2014, there was no significant difference in pre-intervention trends for six out of ten outcome variables (delivery care, ANC, PNC, FP, U5, tetanus vaccination). However, we identified a significant decrease in month-to-month (trend) service provision in the intervention compared with control facilities for GM1, GM2, and OPD. This means that for these services, provision saw faster growth in control than in intervention facilities. We also found an increase in the pre-intervention trend in service provision for immunisation for children aged less than 12 months in intervention facilities, suggesting that provision grew faster in the intervention than in control facilities already before the implementation of PBF. In summary, the analysis therefore indicates that pre-intervention trends were similar for most service provision indicators, reinforcing the robustness and validity of our findings [52].

3.3. Post-PBF intervention comparison

In the post-PBF intervention comparison analyses, we examined the effect of PBF in the intervention compared with control facilities. The model estimates (see Table 4 and 5) indicate that PBF produced modest increases in the trend of service provision for seven outcome variables as well as a modest immediate change in service provision for four outcome variables. In terms of immediate impact, PBF led to a significant increase of 1.3% and 0.7% in the level of service provision for GM1 and GM2 respectively, but led to a significant immediate decrease of 0.6% in service provision for family planning and 0.3% for immunization of children aged less than 12 months. In terms of trend effect, the results showed that PBF led to an increase of 0.4% in the month-to-month provision of service for ANC, 0.2% for PNC, 0.1% for family planning services, 0.4% for GM1, 0.1% for GM2, 0.4% for U5 and 0.2% for OPD. We also found a 0.1% decrease in the month-to-month service provision for immunisation in intervention facilities. PBF did not affect service provision for delivery care and tetanus vaccination at all.

3.4. Post-PBF+ *gratuité* intervention comparison

Following the introduction of *gratuité* alongside PBF, we no longer saw any positive effect of PBF on service provision, except for complete immunisation of children under 12 months where we observed an immediate increase of 0.3%. However, we also observed a decrease of 0.3%, 0.9% and 0.3% for family planning services, GM1 and GM2 respectively. In terms of trend, we found a

0.3%, 0.4% decrease in month-to-month service provision for delivery care and U5 respectively; and an infinitesimal decrease in service provision for family planning services.

4. **Discussion**

Our study makes a unique contribution to the literature in being the first study employing a robust quasi-experimental analytical approach to assess the impact of PBF on the quantity and the mix of service provision while accounting for the introduction of a free healthcare policy. The findings highlight that before the introduction of *gratuité*, PBF produced modest increases over time in the provision of most services, without substantially altering service mix. Specifically, between January 2014 and May 2016, PBF produced modest increases across a wide range of services (ANC, PNC, family planning, GM1, GM2, U5 and curative care), a decrease in only one service (child complete immunisation), and no effect on two services (delivery care and tetanus vaccination for pregnant women). Following the introduction of *gratuité*, PBF did not produce further increases in service provision nor led to any changes in the service mix. After June 2016, service provision for two services targeted by the *gratuité*, namely delivery care and under-five consultations, and for family planning (not targeted by *gratuité*) grew less significantly in PBF compared with non-PBF facilities.

Appraising findings in light of existing literature is difficult given the novelty of the analytical approach we adopted and the wide range of outcome variables included. Comparing our findings with those from the three other impact studies conducted in Burkina Faso, we notice remarkable differences. Discrepancies between our findings and the study by Steenland et al. (2017) are not surprising since the pre-pilot was substantially different from the extended pilot and took place at a time where baseline service provision levels were considerably different from those in 2014 when the extended pilot was launched. In addition, even though Steenland et al. made use of HMIS data, the authors applied a difference-in-differences analytical approach, thereby relying only on two-time points, and worked with data from only three intervention districts. Similarly, differences between our estimates and those produced by the primary IE [25] arise because the analysis we present in this paper relied on multiple data points and could therefore discern the impact of PBF before and after the introduction of the *gratuité*. This underlines that strong routine data collection systems are essential to enable the production of robust impact assessments to complement and, in the case of scarcity of resources, even replace IE based on costly primary data collection [13, 14]

We also identify remarkable differences if we compare our findings with those of the study by Zizien et al. (2018) which is closest to ours. Also addressing the extended PBF pilot, the study detected a positive effect exclusively on the provision of PNC services. The striking difference between the two sets of results may be due to the fact that Zizien et al. relied on a shorter time series (which therefore gave fewer observation points and reflected a shorter intervention period), different control districts (not the ones selected ex-ante for the primary IE), and made use of service provision counts aggregated at the district level rather than disaggregated at facility level as we did. Knowing how the local HMIS operates, we postulate that district-level aggregate service provision counts are more likely to be inaccurate since they are routinely computed by simply equating missing values to zeros, which potentially distorts resulting estimates. Moreover, the analysis by Zizien et al. also stretched into the *gratuité* period, but did not account for it analytically, introducing an additional bias in the generation of PBF effects.

Two interesting factors emerge from our analysis: PBF produced consistent, albeit modest, increases in service provision prior to the introduction of the *gratuité* and, contrary to our initial expectation, it did so without altering the overall service mix. Appraising these two elements jointly suggests that carefully designed PBF programs can effectively increase service provision without necessarily altering the overall service mix. This observation provides evidence against the claim that PBF inevitably introduces gaming, which we define as redirecting efforts towards remunerated services at the expense of non-remunerated aspects of work as well as purposive misreporting[53-55]. While findings from a parallel process evaluation suggest the presence of gaming in selected facilities [56], our quantitative findings clearly indicate that gaming is unlikely to have taken place on a large scale in the Burkinabè context. This observation is consistent with our prior work on the same PBF program, suggesting the existence of widespread misreporting at the facility level, but the absence of systematic misreporting only on selected indicators [57]. The absence of extensive gaming is perhaps a consequence of the fact that PBF in Burkina Faso addressed a broad range of primary healthcare services and relied on unit prices that reflected adequately the relative level of effort implied by provision of the single services. This finding bears important policy implications, as it suggests that PBF may serve as a means of introducing elements of strategic purchasing in a healthcare system [58], but to do so it needs to target a broad range of services and apply an adequate pricing structure [59]. Nonetheless, in line with existing literature [60, 61], the small magnitude of the changes observed raises important questions related to the cost of PBF and calls for further analysis to determine its economic value.

In line with prior literature [7, 21, 62-64], not all services experienced an increase following the introduction of PBF. In particular, contrary to the primary IE [33], we did not identify an effect of PBF on the use of facility-based delivery. While likely due to the different modelling approaches adopted by the two studies, the discrepancy between our findings and the findings from the primary IE remains to be explored and understood through further qualitative inquiry in order to be able to inform prospective policy formulation. At this stage, however, we can speculate that the extended PBF pilot was launched eight years after the introduction of a major user-fee reduction policy targeting specifically delivery services such as Emergency Obstetric and Newborn Care (*Soins Obstétricaux Néonataux d'Urgences*; SONU) [65, 66]. Even prior to the *gratuité*, this policy had allowed Burkina Faso to reach some of the highest rates of assisted deliveries for the African content, with 86.2% of all pregnant women choosing to give birth in a health facility in 2014 [43]. Hence, the trend towards increased facility-based delivery had been upward for several years, offering little opportunities for an intervention to easily produce additional marginal gains.

In line with the primary IE and consistent with observations emerging elsewhere in SSA [63, 67], PBF did not substantially affect the provision of vaccination services, even producing a small decline in the proportion of fully immunised children. This single negative effect detected in the pre-*gratuité* period is surprising, since it contradicts our initial expectation that PBF would be more likely to produce changes in the provision of services that are more susceptible to healthcare providers' behaviour. Given the central role that immunisation plays in public health, understanding the inability of PBF to generate even modest positive changes at this level merits further qualitative inquiry. In line with what was observed in other settings [68], it is possible, but remains to be verified empirically, that structural constraints related to vaccine availability restricted healthcare providers' ability to generate change on this specific service provision indicator. Informal discussions with implementing stakeholders revealed substantial vaccine ruptures in the government supply chain system, particularly towards the end of 2017. It appears that healthcare providers were unable to circumvent such vaccine ruptures since PBF was not accompanied by any measure to enable them to procure drugs outside the standard government supply system.

The third interesting element emerging from our analysis is the fact that PBF did not produce further increases in service provision following the introduction of the *gratuité*. Contrary to what was postulated in our initial theoretical framework, the increases in demand promoted by the *gratuité* through the removal of user fees and convergence of two provider payments on the same sub-set of services did not generate any change in the quantity and mix of services provided by PBF facilities. This inability to generate further increases in service provision is possibly due to a saturation effect, in that in a context characterised by inadequate human resources [69] and overall structural constraints [70], the implementation of *gratuité* alongside PBF probably placed an additional burden on healthcare providers already working at maximum capacity. It is also plausible to assume that by the time the *gratuité* was launched (over two years into PBF implementation), healthcare providers might have been reluctant to make additional efforts to increase service provision, given their experience with poor implementation fidelity, and more specifically with the extreme delays in payment incurred by the PBF intervention **[71, 72]**. Re-iterating the recommendation that has emerged from a review of 32 PBF programs in LMICs [4], our findings also point to the need to conduct further qualitative research to understand how existing health system structures and processes related to PBF implementation interact to shape the program effects (or lack thereof).

Methodological considerations

Next to its value as a sound quasi-experimental investigation into the effects of PBF in Burkina Faso, we acknowledge that our study also suffered from a number of limitations. First, despite overall comparatively good quality of HMIS in Burkina, an earlier study focused on misreporting [57] had identified small discrepancies between PBF verified data and HMIS data. Hence, we cannot exclude potential sources of bias in our analysis. Second, the datasets available for our analyses were not complete, forcing us to drop 221 (out of 1,059) facilities due to a lack of sufficient data to even run multiple imputations. Most of these facilities dropped were recently opened and thus did not have enough observations for the study period. While we trust that the 838 facilities included in the final analysis represent a sufficiently large and representative sample, we cannot fully exclude that they differ in some systematic way from the overall census of facilities in the concerned districts. Unfortunately, we have no means to assess the presence of such potential systematic differences. Nonetheless, the excluded facilities were similarly distributed across intervention and control districts suggesting that data completeness was not directly related to the PBF program. Third, even for the sub-sample of facilities that could be retained for analysis, the datasets contained substantial proportions of missing values for several outcome variables, forcing us to rely on multiple imputations prior to running our models. Albeit clearly representing a second-best option to working with real complete data, in line with the literature on missing data [45, 46, 51, 73], we trust that this procedure produced less biased results than analysing incomplete data. Finally, for four services, namely GM1, GM2, immunisation and curative services, the models identified significant pre-intervention differences in trend, suggesting that for these indicators PBF and non-PBF facilities may not be fully comparable and hence the related coefficients should be interpreted with caution [52].

5. Conclusion

In conclusion, we would like to highlight the two main contributions of our work. First, from a methodological point of view, by relying exclusively on secondary routine data to examine the effect of PBF on the quantity and mix of service provision while accounting for the introduction of a parallel free healthcare policy, our study contributes evidence in favour of the feasibility of implementing such quasi-experimental methods and should serve to encourage other teams working on health financing impact assessments to pursue similar approaches. Second, and more importantly so, our study offers independent evidence that the PBF program produced consistent, yet modest improvements across a wide range of services, not altering service provision mix. This suggests that broadly defined PBF programs can effectively prevent that healthcare providers privilege provision of some services above others. By highlighting how no further gains in terms of service provision were observed after the introduction of the gratuité, our study also points at the importance of considering changing policy environments when evaluating the effects of a specific health financing reform. Our findings invite health policy makers to consider how the health system provision capacity may be capped by structural constraints, so that no further improvements are possible even when two potentially synergetic policies are implemented in parallel to one another. Since this study is based exclusively on quantitative data, we need to acknowledge our inability to explaining the interaction between the two policies, PBF and free healthcare, inevitably leaving open many questions to be addressed by further qualitative research.

References

1. Fritsche, B.R., R. Soeters, and B. Meessen, *Performance-Based Financing Toolkit*. 2014, 1818 H Street NW, Washington DC 20433 International Bank for Reconstruction and Development / The World Bank.
2. Witter, S., et al., *Performance-based financing as a health system reform: mapping the key dimensions for monitoring and evaluation.* BMC Health Serv Res, 2013. **13**: p. 367.
3. Renmans, D., et al., *Performance-based financing: the same is different.* Health Policy and Planning, 2017. **32**(6): p. 860-868.
4. Renmans, D., et al., *Opening the 'black box' of performance-based financing in low- and lower middle-income countries: a review of the literature.* Health Policy Plan, 2016. **31**(9): p. 1297-309.
5. Binyaruka, P., et al., *Effect of Paying for Performance on Utilisation, Quality, and User Costs of Health Services in Tanzania: A Controlled Before and After Study.* PLoS One, 2015. **10**(8): p. e0135013.
6. Friedman, J., et al., *Impact evaluation of Zambia’s health results-based financing pilot project*. 2016, World Bank Group: Washington, D.C. .

7. Bonfrer, I., E. Van de Poel, and E. Van Doorslaer, *The effects of performance incentives on the utilization and quality of maternal and child care in Burundi.* Soc Sci Med, 2014. **123**: p. 96-104.
8. Bonfrer, I., et al., *Introduction of performance-based financing in burundi was associated with improvements in care and quality.* Health Aff (Millwood), 2014. **33**(12): p. 2179-87.
9. Bonfrer, I., et al., *The effects of performance-based financing on the use and quality of health care in Burundi: an impact evaluation.* The Lancet, 2013. **381**: p. S19.
10. Skiles, M.P., et al., *The effect of performance-based financing on illness, care-seeking and treatment among children: an impact evaluation in Rwanda.* BMC Health Serv Res, 2015. **15**: p. 375.
11. Das, A., S.S. Gopalan, and D. Chandramohan, *Effect of pay for performance to improve quality of maternal and child care in low- and middle-income countries: a systematic review.* BMC Public Health, 2016. **16**: p. 321.
12. Hox, J.J. and H.R. Boeije, *Data collection, primary versus secondary.* 2005.
13. Bernal, J.L., S. Cummins, and A. Gasparrini, *Interrupted time series regression for the evaluation of public health interventions: a tutorial.* International journal of epidemiology, 2017. **46**(1): p. 348-355.
14. Lagarde, M., *How to do (or not to do) ... Assessing the impact of a policy change with routine longitudinal data.* Health Policy Plan, 2012. **27**(1): p. 76-83.
15. Nguyen, H.T., et al., *The impact of reducing and eliminating user fees on facility-based delivery: a controlled interrupted time series in Burkina Faso.* Health Policy Plan, 2018. **33**(8): p. 948-956.
16. Zombre, D., M. De Allegri, and V. Ridde, *Immediate and sustained effects of user fee exemption on healthcare utilization among children under five in Burkina Faso: A controlled interrupted time-series analysis.* Soc Sci Med, 2017. **179**: p. 27-35.
17. Steenland, M., et al., *Performance-based financing to increase utilization of maternal health services: Evidence from Burkina Faso.* SSM-population health, 2017. **3**: p. 179-184.
18. Zizien, Z.R., et al., *Contribution of the results-based financing strategy to improving maternal and child health indicators in Burkina Faso.* Int J Health Plann Manage, 2018.
19. Janssen, W., et al., *Improving quality through performance-based financing in district hospitals in Rwanda between 2006 and 2010: a 5-year experience.* Trop Doct, 2015. **45**(1): p. 27-35.
20. Rusa, L., et al., *Performance-based financing for better quality of services in Rwandan health centres: 3-year experience.* Trop Med Int Health, 2009. **14**(7): p. 830-7.
21. Falisse, J.B., et al., *Performance-based financing in the context of selective free health-care: an evaluation of its effects on the use of primary health-care services in Burundi using routine data.* Health Policy Plan, 2015. **30**(10): p. 1251-60.
22. Ministère de la Santé Burkina Faso, *Guide de mise en oeuvre du financement base sur les résultats dans le secteur de la santé.* 2013.
23. Ministère de la Santé du Burkina, F., *Stratégie nationale de gratuité des soins chez la femme enceinte, des accouchements, de la césarienne et des autres SONU*. 2016.
24. Ministère de la Santé du Burkina, F., *Stratégie nationale d'exemption de paiement des soins curatifs pour les enfants de moins de 5 ans*. 2016.
25. De Allegri, M., J. Lohmann, and M. Schleicher, *Results-Based Financing for Health: Impact Evaluation in Burkina Faso*. 2018.
26. World Health Organization. *World Health Report 2015*. 2015 2016 [cited 2016 30/12/2016]; Available from: http://www.who.int/gho/publications/world_health_statistics/EN_WHS2015_Part1.pdf?ua=1.
27. World Health Organization, *country Cooperation Strategy at a glance: Burkina Faso*. 2016.
28. The World Bank. *Burkina Faso*. 2019 [cited 2019 11/07/2019]; Available from: https://www.worldbank.org/en/country/burkinafaso/overview.

29. Burkina Faso Ministère de la santé Unité - Progrès - Justice, *Strategie Nationale de Financement de la Sante Pour la Couverture Sanitaire Universelle 2018 - 2030*. 2018.
30. Munos, M., et al., *Independent evaluation of the rapid scale-up program to reduce under-five mortality in Burkina Faso.* The American journal of tropical medicine and hygiene, 2016. **94**(3): p. 584-595.
31. Burkina Faso Ministere de la Sante secretariat General Unité-Progrès-Justice Direction on Generale des Etudes Et des Statistiques Sectorielles, *Annuaire Statistique 2014*. 2015.
32. De Allegri, M., et al., *Responding to Policy Makers' Evaluation Needs: Combining Experimental and Quasi-Experimental Approaches to Estimate the Impact of Performance Based Financing in Burkina Faso*. In Press.
33. Mwase, T., et al., *Can combining Performance-Based Financing with Equity Measures result in greater equity in utilization of maternal care services? Evidence from a Cluster Randomized Controlled Trial in Burkina Faso.* Submitted.
34. Michielutte, R., et al., *Use of an interrupted time-series design to evaluate a cancer screening program.* Health Educ Res, 2000. **15**(5): p. 615-23.
35. Fretheim, A., et al., *A reanalysis of cluster randomized trials showed interrupted time-series studies were valuable in health system evaluation.* J Clin Epidemiol, 2015. **68**(3): p. 324-33.
36. Shadish, W.R. and T.D. Cook, *The renaissance of field experimentation in evaluating interventions.* Annu Rev Psychol, 2009. **60**: p. 607-29.
37. Kontopantelis, E., et al., *Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis.* Bmj, 2015. **350**: p. h2750.
38. Devkaran, S. and P.N. O'Farrell, *The impact of hospital accreditation on quality measures: an interrupted time series analysis.* BMC Health Serv Res, 2015. **15**: p. 137.
39. Serumaga, B., et al., *Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study.* BMJ, 2011. **342**: p. d108.
40. Jacob, R., et al., *The Validity of the Comparative Interrupted Time Series Design for Evaluating the Effect of School-Level Interventions.* Eval Rev, 2016.
41. Laffont, J.-J. and D. Martimort, *The theory of incentives: the principal-agent model*. 2009: Princeton university press.
42. Sekwat, A., *Principal-agent theory: a framework for improving health care reform in Tennessee.* J Health Hum Serv Adm, 2000. **22**(3): p. 277-91.
43. Ministère de la Santé du Burkina Faso, *Annuaire statistique 2015*. 2016.
44. Burkina Faso Ministere de la Sante Unité - Progrès - Justice, *Manuel de Procedures Descriptives des Modalites de Gestion, de Suivi et de Controle des Mesures de Gratuite des Soins au Profit des Femmes et des Enfants de Moins de cinq (05) ans Vivant au Burkina Faso*. 2018.
45. Dong, Y. and C.Y. Peng, *Principled missing data methods for researchers.* Springerplus, 2013. **2**(1): p. 222.
46. Honaker, J. and G. King, *What to do About Missing Values in Time Series Cross-Section Data.* American Journal of Political Science, 2010. **54**(3): p. 561-581.
47. StataCorp LP, *Stata Multiple Imputation Reference Manual. Release 13*. 2013.
48. Shin, Y., *Time Series Analysis in the Social Sciences*

The Fundamentals. 1 ed. 2017: University of California Press.
49. Linden, A., *Using forecast modelling to evaluate treatment effects in single-group interrupted time series analysis.* J Eval Clin Pract, 2018.
50. Wagner, A.K., et al., *Segmented regression analysis of interrupted time series studies in medication use research.* J Clin Pharm Ther, 2002. **27**(4): p. 299-309.

51. Horton, N.J. and K.P. Kleinman, *Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models.* Am Stat, 2007. **61**(1): p. 79-90.
52. Linden, A., *Conducting Interrupted Time-series Analysis for Single- and Multiple-group Comparisons.* The Stata Journal, 2015. **15**(2): p. 480-500.
53. Conrad, D.A., *The Theory of Value-Based Payment Incentives and Their Application to Health Care.* Health Serv Res, 2015. **50 Suppl 2**: p. 2057-89.
54. Oxman, A.D. and A. Fretheim, *Can paying for results help to achieve the Millennium Development Goals? Overview of the effectiveness of results-based financing.* J Evid Based Med, 2009. **2**(2): p. 70-83.
55. Kalk, A., F.A. Paul, and E. Grabosch, *'Paying for performance' in Rwanda: does it pay off?* Trop Med Int Health, 2010. **15**(2): p. 182-90.
56. Turcotte-Tremblay, A.M., I.A. Gali-Gali, and V. Ridde, *Uncovering the Unintended Consequences of Performance-Based Financing in Burkina Faso.* Forthcoming.
57. Kuunibe, N., et al., *Factors associated with misreporting in performance-based financing in Burkina Faso: Implications for risk-based verification.* Int J Health Plann Manage, 2019.
58. World Health Organization, *Strategic Purchasing for Universal Health Coverage: Unlocking the Potential. GLOBAL MEETING SUMMARY AND KEY MESSAGES*. 2017.
59. Chimhutu, V., I. Lindkvist, and S. Lange, *When incentives work too well: locally implemented pay for performance (P4P) and adverse sanctions towards home birth in Tanzania - a qualitative study.* BMC Health Serv Res, 2014. **14**: p. 23.
60. Turcotte-Tremblay, A.M., et al., *Does performance-based financing increase value for money in low- and middle- income countries? A systematic review.* Health Econ Rev, 2016. **6**(1): p. 30.
61. Chi, Y.-L., et al., *Mind the costs, too: towards better cost-effectiveness analyses of PBF programmes.* BMJ global health, 2018. **3**(5): p. e000994.
62. Rudasingwa, M., R. Soeters, and M. Bossuyt, *The effect of performance-based financial incentives on improving health care provision in Burundi: a controlled cohort study.* Glob J Health Sci, 2014. **7**(3): p. 15-29.
63. Basinga, P., et al., *Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation.* Lancet, 2011. **377**(9775): p. 1421-8.
64. Lannes, L., et al., *Can performance-based financing help reaching the poor with maternal and child health services? The experience of rural Rwanda.* Int J Health Plann Manage, 2016. **31**(3): p. 309-48.
65. Ganaba, R., et al., *The obstetric care subsidy policy in Burkina Faso: what are the effects after five years of implementation? Findings of a complex evaluation.* BMC Pregnancy Childbirth, 2016. **16**: p. 84.
66. Ridde, V., et al., *The national subsidy for deliveries and emergency obstetric care in Burkina Faso.* Health Policy Plan, 2011. **26 Suppl 2**: p. ii30-40.
67. Ashir, G.M., H.V. Doctor, and G.Y. Afenyadu, *Performance based financing and uptake of maternal and child health services in yobe sate, northern Nigeria.* Glob J Health Sci, 2013. **5**(3): p. 34-41.
68. De Allegri, M., et al., *Unraveling PBF effects beyond impact evaluation: results from a qualitative study in Cameroon.* BMJ global health, 2018. **3**(2): p. e000693.
69. Yaya Bocoum, F., et al., *Which incentive package will retain regionalized health personnel in Burkina Faso: a discrete choice experiment.* Hum Resour Health, 2014. **12 Suppl 1**: p. S7.
70. Melberg, A., et al., *'We saw she was in danger, but couldn't do anything': Missed opportunities and health worker disempowerment during birth care in rural Burkina Faso.* BMC Pregnancy and Childbirth, 2016. **16**(1): p. 292.

71. Ridde, V., et al., *Twelve months of implementation of health care performance-based financing in Burkina Faso: A qualitative multiple case study.* Int J Health Plann Manage, 2018. **33**(1): p. e153-e167.
72. Bodson, O., et al., *A study on the implementation fidelity of the performance-based financing policy in Burkina Faso after 12 months.* 2018. **76**(1): p. 4.
73. Pratama, I., et al. *A review of missing values handling methods on time-series data*. in *2016 International Conference on Information Technology Systems and Innovation (ICITSI)*. 2016.

Table 1: Outcome variables (service indicators) and their measurements (numerators and denominators)

Indicator label	*Numerator (per month)*	*Denominator (per year)*[2]
Delivery care	Number of facility based deliveries	Expected deliveries
Tetanus vaccination	Number of second tetanus vaccine doses provided to pregnant women	Expected deliveries
ANC	Antenatal care consultations provided to pregnant women utilizing ANC for the fourth (or more) time in the current pregnancy	Expected deliveries
PNC	Number of postnatal care consultations	Expected deliveries
FP	Number of family planning services provided to women of reproductive age	Women 15-49 years old
GM1	Number of children aged <12 months provided with growth monitoring services	Population aged <12 months
GM2	Number of children aged 12 to 23 months provided with growth monitoring services	Population < 5 years[3]
Immunization	Number of children aged < 12 months provided with complete immunization	Population < 12 months
U5	Curative care consultations provided to children aged < 5 years	Population < 5 years
OPD	Curative care consultations provided to people aged 5 years and above	Total population

[2] Using annual data as denominator resulted in lower values (for proportion). Since it is not likely that population figures vary per month, we preferred using annual projections for all indicators for consistency purposes.

[3] We did not get the specific match for the HMIS data.

Table 2: Selection of model functional form

Outcomes	Linear		Quadratic		Semi-log	
	AIC[4]	BIC[5]	AIC	BIC	AIC	BIC
ANC	-655.7	-649.2	**-656.0**[6]	**-649.6**	-146.0	-139.6
Tetanus vaccination	-1326.0	-1319.6	**-1326.4**	**-1319.9**	-121.8	-115.4
Delivery care	**-953.1**	**-946.6**	-952.9	-946.5	19.8	26.3
PNC	-994.1	-987.6	**-994.1**	**-987.6**	-116.7	-110.2
Immunization	-1624.2	-1617.8	**-1624.7**	**-1618.3**	-216.3	-209.8
FP	**-1362.3**	**-1355.9**	-1362.1	-1355.6	-35.9	-29.5
GM1	-1034.0	-1027.6	**-1034.6**	**-1028.1**	-167.4	-160.9
GM2	**-1278.8**	**-1272.3**	-1278.4	-1271.9	-100.4	-93.9
U5	-705.8	-699.4	**-705.9**	**-699.5**	-109.4	-102.9
OPD	-1113.3	-1106.9	**-1114.5**	**-1108.1**	-39.9	-33.5

[4] Akaike information criterion
[5] Bayesian information criterion
[6] Selected model

Table 3: Average monthly rate (proportions) of healthcare service provision

Outcome	PBF facilities				Non-PBF facilities			
	Mean	**Std. Dev**	**Min**	**Max**	**Mean**	**Std. Dev**	**Min**	**Max**
Pre-intervention period (January 2013 to December 2013 - 12 months)								
ANC	0.241	0.011	0.222	0.256	0.245	0.006	0.232	0.258
Tetanus vaccination	0.034	0.002	0.029	0.038	0.041	0.002	0.039	0.044
Delivery care	0.067	0.004	0.058	0.075	0.069	0.006	0.057	0.078
PNC	0.087	0.004	0.081	0.093	0.089	0.004	0.085	0.095
Immunization	0.050	0.001	0.049	0.052	0.048	0.001	0.047	0.051
FP	0.025	0.002	0.022	0.029	0.025	0.001	0.024	0.028
GM1	0.089	0.003	0.085	0.094	0.042	0.002	0.037	0.045
GM2	0.041	0.002	0.037	0.045	0.097	0.006	0.087	0.106
U5	0.195	0.011	0.172	0.214	0.198	0.006	0.184	0.208
OPD	0.045	0.001	0.042	0.048	0.051	0.002	0.045	0.054
Implementation of PBF only (January 2014 to May 2016 – 29 months)								
ANC	0.243	0.019	0.213	0.279	0.241	0.016	0.220	0.274
Tetanus vaccination	0.038	0.002	0.033	0.044	0.043	0.002	0.039	0.046
Delivery care	0.075	0.004	0.064	0.082	0.072	0.009	0.058	0.089
PNC	0.089	0.007	0.079	0.104	0.092	0.006	0.082	0.101
Immunization	0.049	0.001	0.045	0.051	0.049	0.002	0.046	0.052
FP	0.026	0.003	0.021	0.032	0.027	0.003	0.023	0.032
GM1	0.091	0.007	0.078	0.107	0.098	0.006	0.089	0.110
GM2	0.043	0.003	0.037	0.047	0.042	0.003	0.036	0.049
U5	0.191	0.018	0.163	0.223	0.200	0.016	0.177	0.237
OPD	0.049	0.005	0.042	0.057	0.053	0.004	0.046	0.063
Implementation of gratuité alongside PBF (June 2016 to September 2017 – 16 months)								
ANC	0.264	0.012	0.237	0.285	0.243	0.016	0.217	0.264
Tetanus vaccination	0.039	0.002	0.035	0.041	0.043	0.002	0.039	0.045
Delivery care	0.073	0.005	0.065	0.078	0.071	0.007	0.056	0.081
PNC	0.092	0.005	0.082	0.099	0.087	0.006	0.075	0.095
Immunization	0.048	0.001	0.047	0.049	0.049	0.001	0.048	0.052
FP	0.028	0.001	0.025	0.030	0.025	0.002	0.022	0.028
GM1	0.099	0.003	0.095	0.104	0.097	0.006	0.089	0.107

GM2	0.047	0.003	0.041	0.052	0.039	0.002	0.038	0.043
U5	0.215	0.009	0.199	0.229	0.196	0.019	0.165	0.221
OPD	0.058	0.005	0.051	0.067	0.053	0.005	0.046	0.062
Study period (January 2013 to September 2017 - 57 months)								
ANC	0.248	0.018	0.213	0.284	0.242	0.015	0.216	0.274
Tetanus vaccination	0.037	0.003	0.029	0.044	0.042	0.002	0.039	0.047
Delivery care	0.072	0.005	0.058	0.081	0.071	0.008	0.056	0.089
PNC	0.089	0.006	0.079	0.104	0.089	0.006	0.075	0.102
Immunization	0.048	0.001	0.046	0.053	0.049	0.002	0.045	0.052
FP	0.027	0.003	0.021	0.032	0.026	0.002	0.022	0.032
GM1	0.092	0.006	0.078	0.106	0.098	0.006	0.087	0.110
GM2	0.044	0.004	0.036	0.052	0.041	0.002	0.036	0.049
U5	0.198	0.018	0.163	0.229	0.199	0.015	0.165	0.237
OPD	0.051	0.006	0.042	0.067	0.053	0.004	0.045	0.063

Table 4: Maternal care services *[GLS based estimates (autocorrelation adjusted model)]*

Maternal care services	**Delivery care**	**Tetanus vaccination**	**ANC**	**PNC**	**FP**
Baseline situation					
Level *(control group)* (θ_0)	0.069***	0.042***	0.237***	0.089***	0.025***
Trend *(control group)* (θ_1)	0.001	-3.09e-04***	0.002	1.39e-04	1.36e-04
Difference in level (θ_4)	-0.008	-0.008***	0.001	-3.73e-04	0.003*
Difference in trend (θ_5)	4.013e-04	1.68e-04	-0.002	-7.38e-04	-3.81e-04*
Effect of original PBF					
Level change *(control group)* (θ_2)	0.008	0.005***	-0.008	0.005*	0.005***
Change in trend *(control group)* (θ_3)	-1.61e-04	-4.07e-04	-0.002	-3.04e-04	-122e-04**
Difference in level (θ_6)	-0.004	-0.001	-0.002	-0.002	-0.006***
Difference in trend (θ_7)	2.64e-04	6.61e-06	0.004**	0.002**	0.001***
Effect of PBF+gratuité					
Level change *(control group)* (θ_8)	0.002	-0.001	0.002	-0.001	0.001
Change in trend *(control group)* (θ_9)	0.003**	-3.51e-04	0.003	0.001	0.001**
Difference in level (θ_{10})	0.005	-0.002	-0.019	-0.005	-0.003***
Difference in trend (θ_{11})	-0.003***	-2.803e-04	-0.003	-0.001	-4.66e-04**
Quadratic term	-2.06e-05	1.18e-05**	-1.03e-05	-3.67e-06	-5.09e-06
Autocorrelation					
rho	-0.198	0.616	0.697	0.711	0.728
DW statistic (unadjusted)	2.365	0.980	0.777	0.879	0.614
DW statistic (adjusted)	2.025	1.938	2.064	1.912	1.817

*** = significant at 1 percent level; ** = significant at 5 percent level; * =significant at 10 percent level

Table 5: curative and child-care services *[GLS based estimates (autocorrelation adjusted model)]*

Child and curative care services	**GM1**	**GM2**	**Immunization**	**U5**	**OPD**
Baseline situation					
Level *(control group)* (θ_0)	0.088***	0.039***	0.050***	0.191***	0.046***
Trend *(control group)* (θ_1)	0.002***	0.001**	-1.25e-04*	0.002	0.001***
Difference in level (θ_4)	0.004*	0.003	-0.001	2.84e-04	3.81e-04
Difference in trend (θ_5)	-0.003***	-0.001**	4.41e-04***	-0.002	-0.001***
Effect of original PBF					
Level change *(control group)* (θ_2)	-0.006**	-0.001	-0.001	0.004	0.001
Change in trend *(control group)* (θ_3)	-0.003***	-0.001	0.001***	-0.003	-0.002***
Difference in level (θ_6)	0.013***	0.007***	-0.003***	-0.158	0.003
Difference in trend (θ_7)	0.004***	0.001**	-0.001***	0.004**	0.002***
Effect of PBF+gratuité					
Level change *(control group)* (θ_8)	0.003	0.001	-0.002***	-0.005	-0.001
Change in trend *(control group)* (θ_9)	-1.76e-04	2.0e-04	4.61e-04**	0.004*	0.001
Difference in level (θ_{10})	-0.009***	-0.003***	0.003***	-0.002	-0.001
Difference in trend (θ_{11})	-4.37e-04	0.001	5.81e-05	-0.004**	-0.001
Quadratic term	1.45e-05	-3.24e-07	-9.19e-06***	-2.37e-07	6.95e-06
Autocorrelation					
rho	0.775	0.709	0.597	0.694	0.650
DW statistic (unadjusted)	0.647	0.729	0.875	0.782	0.798
DW statistic (adjusted)	1.852	1.892	1.784	1.952	1.931

*** = significant at 1 percent level; ** = significant at 5 percent level; * =significant at 10 percent level

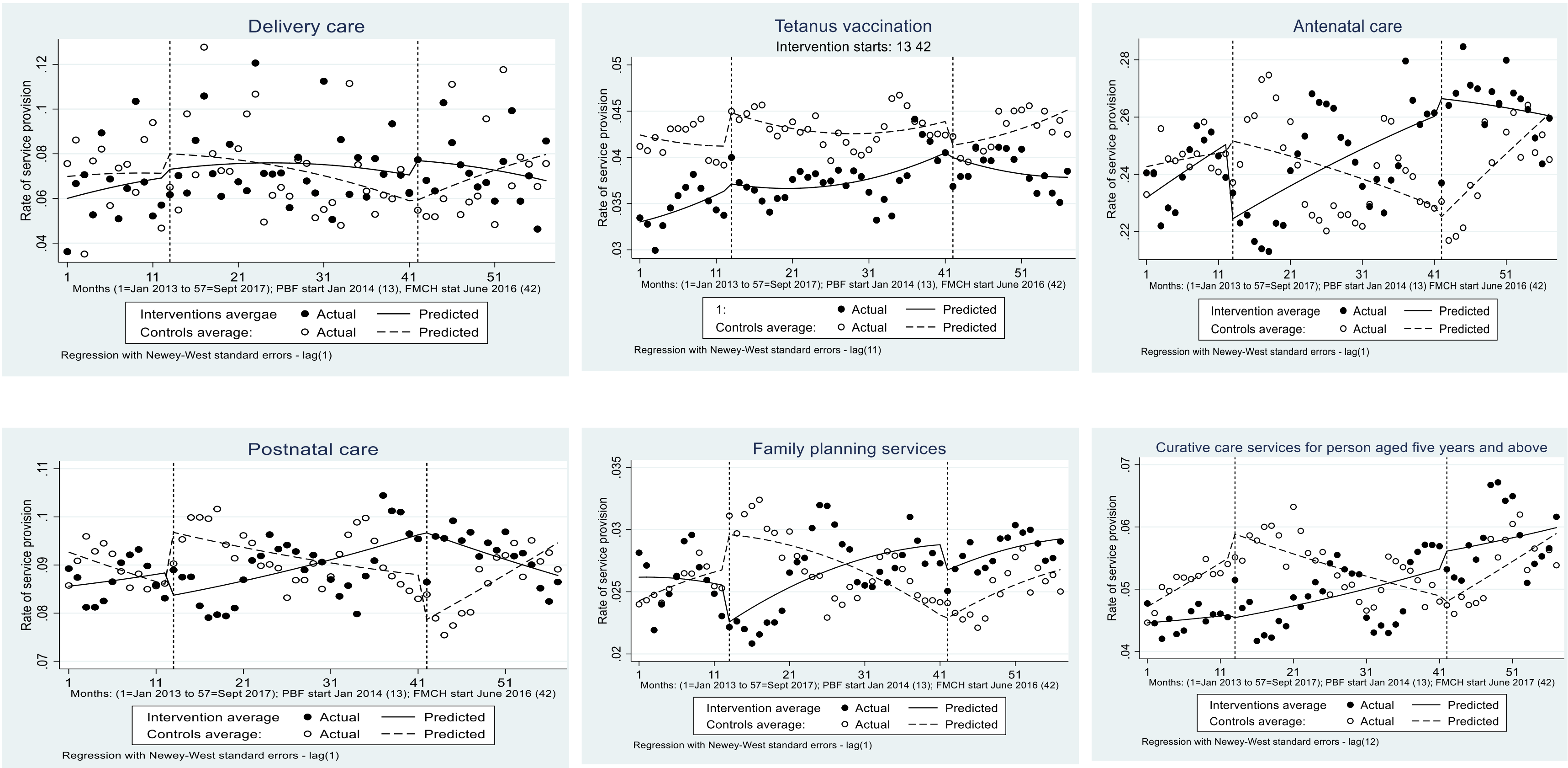

Figure 1: Interrupted time series graphs for maternal and curative care services

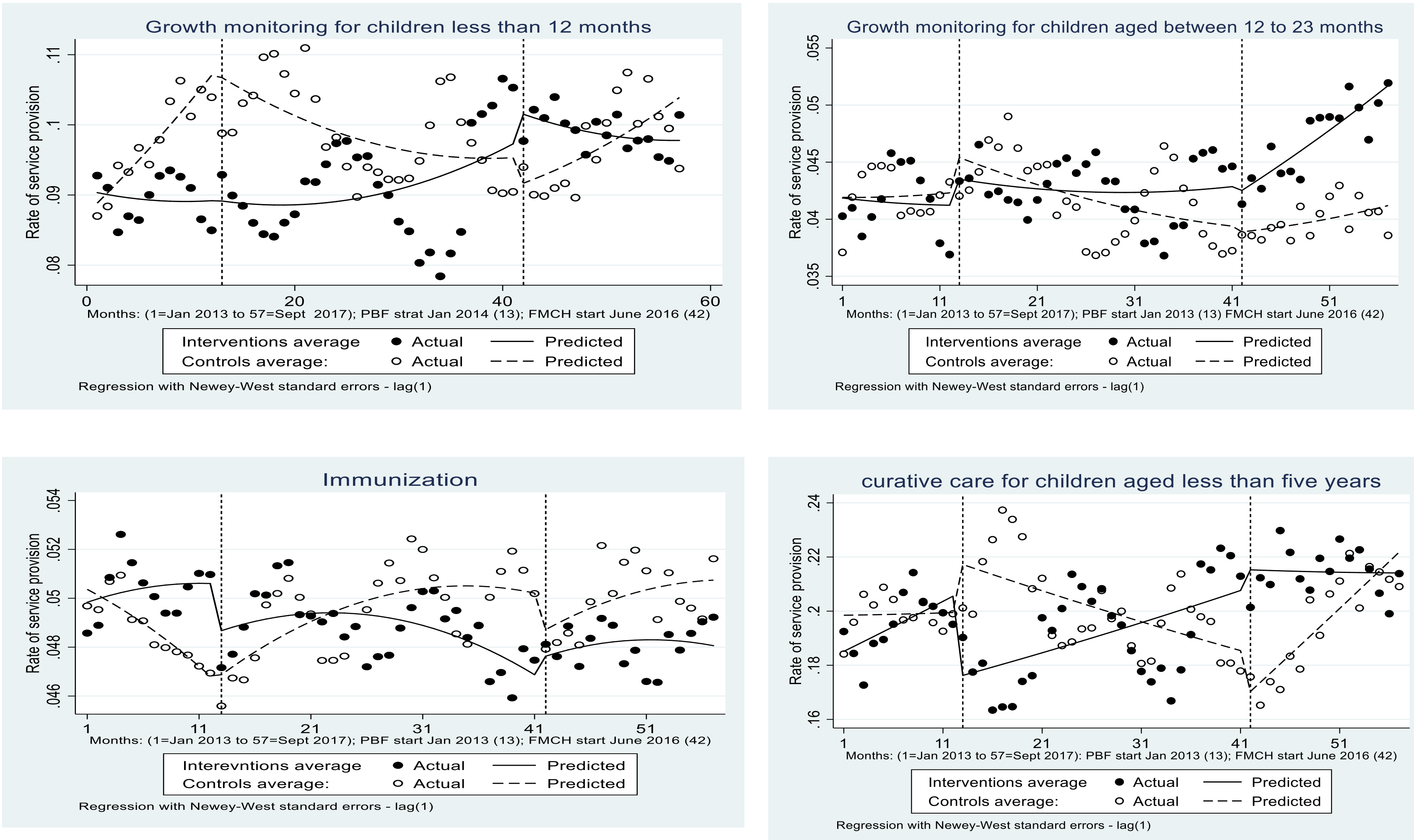

Figure 2: Interrupted time series graphs for childcare services

Appendix A: List of quantity indicators included in PBF design

No.	*Indicator*	*Basic price*[7]
1	Number of new patients age 5 or older in curative consultation	100
2	Number of new patients under the age of 5 in curative consultation	150
3	Number of days of hospitalization	250
4	Number of counter-references received	1010
5	Number of children fully vaccinated	300
6	Number of pregnant women who have received two or more doses of tetanus vaccine	250
7	Number of pregnant women (new and repeat visits) in antenatal care consultation	400
8	Number of women in postnatal consultation (6-8 days and 6-8 weeks post-delivery)	500
9	Number of deliveries performed	1510
10	Number of women (new and repeat visits) in family planning consultation using oral or injectable contraceptives	605
11	Number of women (new and repeat visits) in family planning consultation using long-term methods (IUD or implant)	1210
12	Number of new patients aged 0-11 months in growth monitoring consultation	100
13	Number of patients aged 12-23 months in growth monitoring consultation	250
14	Number of children aged 6-59 months treated for moderate acute malnutrition	300
15	Number of children aged 6-59 months treated for severe acute malnutrition without complications (SAM)	600
16	Number of home visits effected	3000
17	Number of clients having benefitted from voluntary HIV testing and counselling (excluding pregnant women tested in the context of PMTCT)	500
18	Number of pregnant women having benefitted from voluntary HIV testing and counselling in the context of PMTCT	500
19	Number of HIV-positive mothers having benefitted from complete prophylactic anti-retroviral treatment	2500
20	Number of new-borns to HIV-positive mothers treated	3000
21	Number of people living with HIV under anti-retroviral treatment	1000
22	Number of pulmonary tuberculosis cases (new and relapse) detected	6000
23	Number of tuberculosis cases (all types) treated and declared cured or treatment terminated	8500

[7] Burkina Faso CFA franc

Appendix B: Maternal care services [GLS based estimates (autocorrelation adjusted model)]

Maternal care services	**Delivery**	**Tetanus**	**Antenatal**	**Postnatal**	**Family planning**
Baseline situation					
Level *(control group)* (θ_0)	0.069***	0.052***	0.239***	0.064***	0.025***
Trend *(control group)* (θ_1)	0.001	-1.94e-04	1.03e-04	0.002***	-2.79e-04
Difference in level (θ_4)	0.004	-0.007***	0.013	-0.001	0.003**
Difference in trend (θ_5)	1.49e-04	-3.48e-04	-0.001	4.39e-04	1.84e-04*
Effect of original PBF					
Level change *(control group)* (θ_2)	5.03e-04	6.30e-04	0.025***	-0.004***	0.001***
Change in trend *(control group)* (θ_3)	-0.001	-7.41e-04**	-0.007***	2.34e-04	-2.26e-04
Difference in level (θ_6)	-0.001	0.003***	0.003	-0.004**	-4.58e-04
Difference in trend (θ_7)	-1.22e-04	1.41e-04	0.002	4.31e-04	-2.81e-04**
Effect of PBF+gratuité					
Level change *(control group)* (θ_8)	-0.003	0.001**	-0.002	0.001	-0.001*
Change in trend *(control group)* (θ_9)	0.001	1.33e-04	-0.002	0.001	-0.001***
Difference in level (θ_{10})	0.001	-0.001*	0.002	5.59e-04	3.75e-04
Difference in trend (θ_{11})	0.001	-5.77e-04	-0.001	-0.001	8.44e-04
Quadratic term	2.95e-08	7.89e-06	1.13e-04	-2.49e-04**	7.49e-06
Autocorrelation					
rho	0.317	0864	0.704	0.775	0.782
DW statistic (original)	1.735	0.791	0.704	0.555	0.479
DW statistic (transformed)	1.578	1.534	1.701	1.074	1.264

Appendix C: Curative and childcare services [GLS based estimates (autocorrelation adjusted model)]

Maternal care services	**GM1**	**GM2**	**Immunisation**	**U5**	**Curative care**
Baseline situation					
Level *(control group)* (θ_0)	0.108***	0.038***	0.021***	0.142***	0.039***
Trend *(control group)* (θ_1)	-0.002***	-4.0e-04	-2.88e-04	0.005*	3.99e-04
Difference in level (θ_4)	0.003	0.001	0.002**	0.003	-0.001
Difference in trend (θ_5)	-1.65e-06	4.22e-04**	-7.46e-04	0.002	2.20e-04
Effect of original PBF					
Level change *(control group)* (θ_2)	0.027***	3.43e-04	-1.34e-04	-0.065***	0.001
Change in trend *(control group)* (θ_3)	-0.001	-0.001***	-0.001**	-0.001	1.76e-04
Difference in level (θ_6)	-0.001	-0.002***	0.001**	-0.015	3.11e-04
Difference in trend (θ_7)	1.65e-04	-1.72e-04	-1.43e-04	-0.002	-5.99e-04
Effect of PBF+gratuité					
Level change *(control group)* (θ_8)	1.10e-04	-7.19e-04	-0.002***	0.046***	0.005***
Change in trend *(control group)* (θ_9)	-3.11e-04	-0.001***	-0.001**	0.007	0.001
Difference in level (θ_{10})	-0.004	-4.86e-04	-0.001***	0.002	2.88e-04
Difference in trend (θ_{11})	-0.001	-0.001	1.15e-04	0.001	-3.11e-04
Quadratic term	2.74e-04	2.34e-04***	2.71e-04***	-4.79e-04	-6.31e-06
Autocorrelation					
rho	0.858	0.929	0.961	0.844	0.844
DW statistic (original)	0.339	0.431	0.414	0.494	0.406
DW statistic (transformed)	0.793	1.124	0.751	0.675	0.749