# Temi di Discussione
(Working Papers)

Dif-in-dif estimators of multiplicative treatment effects

by Emanuele Ciani and Paul Fisher

BANCA D'ITALIA

EUROSISTEMA

# Temi di discussione

(Working papers)

Dif-in-dif estimators of multiplicative treatment effects

by Emanuele Ciani and Paul Fisher

*The purpose of the* Temi di discussione *series is to promote the circulation of working papers prepared within the Bank of Italy or presented in Bank seminars by outside economists with the aim of stimulating comments and suggestions.*

*The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.*

# DIF-IN-DIF ESTIMATORS OF MULTIPLICATIVE TREATMENT EFFECTS

by Emanuele Ciani* and Paul Fisher**

## Abstract

We consider a difference-in-differences setting with a continuous outcome, such as wages or expenditure. The standard practice is to take the logarithm of the outcome and then interpret the results as an approximation of the multiplicative treatment effect on the original outcome. We argue that a researcher should instead focus on the original outcome when discussing causal inference. Furthermore, it is preferable to use a non-linear estimator, because running OLS on the log-linearised model might confound distributional and mean changes. We illustrate the argument with an original empirical analysis of the impact of the UK Educational Maintenance Allowance on households' expenditure, and with a simulation exercise.

## Contents

_____

* Bank of Italy, Regional Economic Research Division, Florence Branch.

** Institute for Social and Economic Research, University of Essex.

# 1 Introduction[1]

In applied empirical research, it is common to replace continuous outcomes, such as earnings or expenditure, with their logarithm. Often, the choice is motivated by distributional features, like skewness, that a researcher may want to consider. In the difference-in-differences (dif-in-dif) setting, the desire to give a causal interpretation to the estimates complicates the choice. The model the researcher has in mind is usually one with multiplicative effects, which are linearised taking logs. If this is the case, the assumptions needed for causal inference refer to the non-transformed model. In general, this is not explicitly discussed.

To explore the attention received by this issue in the dif-in-dif literature, we reviewed papers published in one top journal with an empirical focus, the Quarterly Journal of Economics, between 2001-2011. A table with complete references is available in Appendix A. In total, 25 papers using a dif-in-dif estimator with continuous outcomes were found. In 9 cases, the outcome is not transformed and an additive model is estimated. We found 16 papers in which at least one outcome is expressed in logarithmic form. The variables most commonly log-transformed are earnings and productivity, followed by a group of other monetary quantities including expenditure, land value, exports and loans. In only 5 out of 16 cases is an explicit reason for the log-transformation given (Manning (1998) lays out common reasons for taking the logarithm of a dependent variable). For example, Nunn and Qian (2011) refer to concerns about skewness in the dependent variable, whereas DellaVigna and Kaplan (2007) state that they wish to account for percentage changes in the control variables. In general, no discussion of the impact of the log transformation on the

causal interpretation is given. Only Finkelstein (2007) states that the OLS estimates for the log of the dependent variable relate to $E(ln(y|x))$, and not $ln(E(y|x))$. To provide estimates of $ln(E(y|x))$, Finkelstein (2007) estimates a generalised linear model (GLM) with log links.

Previous theoretical literature on non-linear dif-in-dif mostly focused on the interpretation of the interaction effect. Mullahy (1999) discussed the case of a log-linearised exponential model. Ai and Norton (2003) showed that in non-linear models the marginal effect of the interaction term is not directly related to its coefficient in the linear index. However, Puhani (2012) recently argued that their way of calculating the marginal effect is not the correct one for the dif-in-dif case. A separate stream of research, not directly related to dif-in-dif, focused on the estimation of exponential models (Mullahy, 1997; Manning, 1998; Manning and Mullahy, 2001; Ai and Norton, 2008). Santos Silva and Tenreyro (2006) showed that the OLS estimator of the log-linearised model may not be consistent for the parameter of interest. Blackburn (2007) discussed how to estimate wage differentials without using logarithms.

In this paper we attempt to reconcile the two streams of research for the dif-in-dif case. Our main aim and contribution is to recollect in a unified setting a number of results that are scattered in the literature, in order to provide the practitioner with a clear guide on the choice of modelling and estimation. Using a potential outcome framework, we reinterpret previous findings to argue that the choice between a multiplicative and an additive model is fundamental to the causal interpretation of the estimands. This choice should be taken before deciding whether or not to take logs, which should be understood as an estimation strategy rather than a matter of model specification.[2]

---

[2]Bertrand et al. (2004) discussed how serial correlation may severely bias inference in dif-in-dif, because conventional standard errors are likely to underestimate the true standard deviation. We do not discuss how to account for this problem in exponential models. However, the main example throughout their paper has $log(wage)$ as the dependent variable, so that the problems discussed here also apply in their context. For instance, one of the possible solutions that they suggested is to collapse the data over the pre-treatment and post-treatment period. Given that this implies

Specifically, we point out in Section 2 that the choice between an exponential or a level model is essentially related to the common trends assumption. Differently, whether the treatment effect is multiplicative or additive does not make a large difference, at least from an $ex-post$ evaluation perspective. Although this follows from different results available in the literature, we could not find a reference that made this important point explicit. In terms of estimation, building on Santos Silva and Tenreyro (2006), but focusing on the dif-in-dif case, we show that using OLS on the log-linearised model may give biased estimates of the true multiplicative effect. This problem can arise if the treatment causes not only a shift in the mean, but also other distributional changes, for instance an increase in the variance for the treated group.

Fortunately, different authors (Mullahy, 1997; Santos Silva and Tenreyro, 2006; Blackburn, 2007) have pointed out that to estimate a multiplicative effect there is no need to log-linearise, because a simple and robust non-linear estimator (Poisson Pseudo-Maximum Likelihood) is available. Although Gregg et al. (2006) noted that it is possible to recover a percentage treatment effect from linear OLS in levels, we point out that one cannot give a causal interpretation to both the additive and multiplicative model. We also correct their calculation in order to properly account for a multiplicative time trend.

We finally show that, in the case of heterogeneous effects, the exponential dif-in-dif model with a conditional mean assumption does not identify the average multiplicative effect, but rather the multiplicative effect on the average. Moreover, the necessary conditions for the latter to be consistently estimated using log-linearisation are less likely to hold with heterogeneous effects. To the best of our knowledge, these two results were not discussed in the dif-in-dif literature, although they are partially related to a comment by Angrist (2001).

In Section 3 we present an original applied example. We study the impact on

---

averaging logs, further research might try to understand whether it introduces a different source of bias.

households' expenditure of the introduction of the Educational Maintenance Allowance (EMA) in the UK. In Section 4 we present a simulation to illustrate our main arguments. Section 5 concludes and summarises the discussion in terms of a guideline for practitioners.

# 2   Model specification and inference

A practitioner willing to estimate a dif-in-dif model with a continuous outcome, such as wages or expenditure, usually faces three main decisions:

1. Shall I model the time trend in additive or in multiplicative form? And shall I report the treatment effect as a difference in levels, or as a percentage change?

2. How can I estimate the multiplicative model? Shall I take logs?

3. What kind of average effect is identified by my model?

We argue that these points should be addressed independently, in order to correctly separate model specification from estimation. The next three subsections are dedicated to these issues.

## 2.1   Multiplicative or additive effects?

In this section we compare an additive dif-in-dif model with one with multiplicative effects. Firstly, we highlight that the key difference is the way in which the time trend is specified. Secondly, we show that, in practice, the two models are related, but that one cannot give a causal interpretation to both.

We start with the simplest, though quite popular dif-in-dif setting, involving two groups ($g \in \{control, \ treated\}$) and two time periods ($t \in \{pre,post\}$), with only one group actually receiving the treatment in the second period. In this paper, we analyse the case of a continuous outcome $y$, such as earnings or consumption.

Several assumptions are required in order to identify the causal effect of the treatment. We draw attention on those related to the functional form. These depend on which feature of the distribution of $y$ we are interested in. Here we focus on the expected value, which is usually the target in program evaluation using dif-in-dif.[3]

First, we specify a model for the expected value of $y$ when non treated ($y_{0igt}$), conditional on $g$ and $t$. The second step is to assume how the expected value of the potential outcome when treated ($y_{1igt}$) is related with the expected $y_{0igt}$. The standard dif-in-dif in levels starts from (Angrist and Pischke, 2009):

$$E\left[y_{1igt}|g,t\right] = E\left[y_{0igt}|g,t\right] + \delta^* = \mu_g^* + \lambda_t^* + \delta^*. \tag{1}$$

where we combine an additive common trends assumption with an additive treatment effect. The superscript $^*$ is used to differentiate the model in levels from the multiplicative one. Note also that receiving the treatment, with a potential outcome $y_{1igt}$, does not coincide with being in the treated group, because in the first period all individuals go untreated.

Differently, one might specify an exponential model

$$E\left[y_{0igt}|g,t\right] = exp\left(\mu_g + \lambda_t\right) \tag{2}$$

where the assumption of common trends is in multiplicative form.[4] Over time, the outcome in the absence of treatment would increase by the same percentage in both

---

[3]One important point to highlight is that in this paper we focus only on the average outcomes. Athey and Imbens (2006) proposed instead a generalised dif-in-dif model that gives a structural interpretation to all differential changes in the distribution of the outcome $y$ over time. Their assumptions on the model of $y$ would therefore be valid for any $f(y)$, where $f(\cdot)$ is a strictly monotone transformation (such as log). Differently, in this paper we give a structural interpretation only to changes in the expected value. We ignore higher moments of the distribution of $y$, which are allowed to change either as a consequence of treatment or time. As noted by Athey and Imbens (2006, pg. 435-436), this approach focused on the conditional mean is not nested in their model, unless one assumes that all individual shocks are statistically independent from group and time.

[4]See Mullahy (1997) for a discussion of IV estimation of an exponential model.

groups. Now we can assume a proportional treatment effect:

$$\frac{E\left[y_{1igt}|g,t\right] - E\left[y_{0igt}|g,t\right]}{E\left[y_{0igt}|g,t\right]} = exp\left(\delta\right) - 1. \tag{3}$$

where $\delta$ is a parameter on the linear index of the exponential model. The multiplicative model for $E\left[y_{1igt}|g,t\right]$ is therefore $exp(\mu_g + \lambda_t + \delta)$. Intuitively, the total percentage change in the expected outcome of the treated group is given by the composition of a percentage change due to time (call it *%time*) and the percentage effect of the treatment (call it *%effect*), so that *(1+%change) = (1+%time) × (1+%effect)*. Differently, for the control group *(1+%change) = (1+%time)*.

To be precise, the key difference between the exponential model and the linear one is in the common trends assumption. The choice of a multiplicative or additive treatment effect plays a less important role. If we are only interested in the *ex-post* evaluation problem, in the spirit of DiNardo and Lee (2011), we may just want to understand which share of the treated-control difference should be attributed to the treatment. However, with multiplicative time trends we still need the counterfactual to be specified as in eq. (2), otherwise we would confound time and treatment effects.

To clarify, Figure 1 is generated with an exponential model as in eq. (2) and (3). In this case the treated group starts from a lower position. Given the multiplicative trend, in the absence of the treatment the increase in this group over time would be smaller in absolute value. Therefore a standard dif-in-dif in levels would underestimate the share of the change that has to be attributed to the treatment. Differently, once we account correctly for the multiplicative time trend, it does not matter whether we express the treatment effect as a percentage difference or as a level difference. Indeed, the former is the fraction on the left hand side of eq. (3), while the latter is simply its numerator. Nevertheless, once the time trend is in multiplicative form, having a multiplicative treatment effect leads to an exponential model, which is clearer and easier to estimate.[5]

---

[5]The situation is different if we are willing to predict how the policy will affect future outcomes.
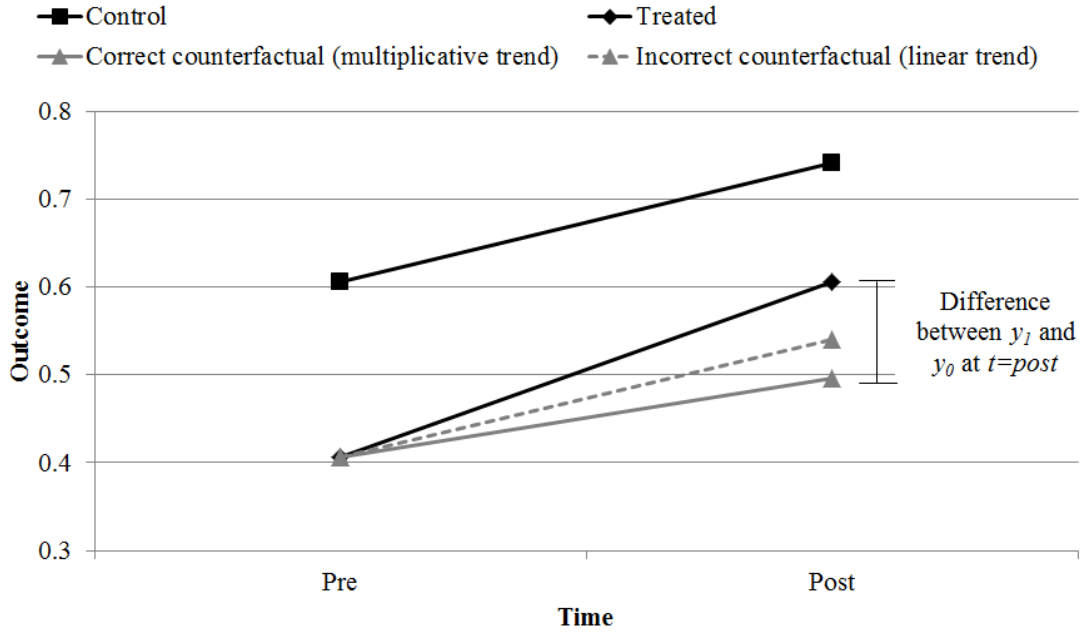
Figure 1: An example of a dif-in-dif setting with multiplicative time trend. The figure shows the expected value of the actual and counterfactual outcome for the different groups according to the exponential model in eq. (2) and (3) with $\mu_{control} = -0.3$; $\mu_{treated} = -0.7$; $\lambda_{pre} = -0.2$; $\lambda_{post} = 0$; $\delta = 0.2$.

Suppose now that we choose a multiplicative model for $y_{1igt}$ and $y_{0igt}$ and we want to understand which would be the correct specification for the observed outcome $y_{it}$. Define the dummies $treated_{it}$ for the treatment group and $post_{it}$ for the second period. The particular data structure leads to an exponential model for observed outcomes

$$E\left[y_{it}|treated_{it}, post_{it}\right] =$$

$$exp\left(\beta_0 + \beta_1 treated_{it} + \beta_2 post_{it} + \delta treated_{it} \times post_{it}\right) \quad (4)$$

$$\beta_0 \equiv \mu_{control} + \lambda_{pre}; \ \beta_1 \equiv \mu_{treated} - \mu_{control}; \ \beta_2 \equiv \lambda_{post} - \lambda_{pre}. \quad (5)$$

In this case the coefficient on $treated_{it} \times post_{it}$ has a meaningful interpretation, because it is directly related to the treatment effect. Indeed, $exp\left(\delta\right)$ is a ratio of

---

If we believe that the treatment is likely to have the same proportional effect in other time periods, then it should be presented in percentage form.

ratios (ROR), as highlighted by Mullahy (1999) and Buis (2010):

$$exp\left(\delta\right) = \frac{E\left[y_{it}|treated_{it} = 1, post_{it} = 1\right]}{E\left[y_{it}|treated_{it} = 1, post_{it} = 0\right]} \bigg/ \frac{E\left[y_{it}|treated_{it} = 0, post_{it} = 1\right]}{E\left[y_{it}|treated_{it} = 0, post_{it} = 0\right]}. \quad (6)$$

Differently, one could follow the well-known suggestion from Ai and Norton (2003) and calculate the interaction effect as the cross difference (Mullahy, 1999, pg. 7):

$$\frac{\Delta^2 E\left[y_{it}|treated_{it}, post_{it}\right]}{\Delta treated_{it}\Delta post_{it}} =$$

$$\left[exp\left(\beta_0 + \beta_1 + \beta_2 + \delta\right) - exp\left(\beta_0 + \beta_1\right)\right] - \left[exp\left(\beta_0 + \beta_2\right) - exp\left(\beta_0\right)\right] \quad (7)$$

which is equal to the change *in levels* of the average for the treated group minus the change *in levels* of the average for the control group. However, following the previous discussion (see Figure 1), this quantity does not properly account for the exponential time trend. Therefore, in a multiplicative model the causal parameter of interest is recovered by the ROR and not by the cross-difference.[6]

In this simplified setting, the exponential model from eq. (4) is actually related to the parameters of a standard dif-in-dif linear regression. Even if the true potential outcomes model is in multiplicative form, the conditional expectation of the observed outcome $y_{it}$ can also be correctly specified as linear:

$$E\left[y_{it}|treated_{it}, post_{it}\right] = \gamma_0 + \gamma_1 treated_{it} + \gamma_2 post_{it} + \tau treated_{it} \times post_{it}. \quad (8)$$

The reason is that the model is saturated: the four parameters fit perfectly the four averages given by the combination of $treated_{it}$ and $post_{it}$. Indeed, the exponential

---

[6]This point is related to the more general comment by Puhani (2012) (also reprised by Karaca-Mandic et al., 2012) who argue that, in any non-linear dif-in-dif model with an index structure and a strictly monotonic transformation function, the treatment effect is not equal to the cross-difference of the observed outcome. In this case, therefore, the well-known suggestion made by Ai and Norton (2003) about how to calculate the cross-difference in non-linear models may be misleading, as it would lead to calculate (7) rather than (6).

model from (4) is just a reparametrisation of the linear one, with

$$exp\left(\delta\right) - 1 = \frac{\left(\gamma_0 + \gamma_1 + \gamma_2 + \tau\right)/\left(\gamma_0 + \gamma_1\right)}{\left(\gamma_0 + \gamma_2\right)/\gamma_0} - 1. \tag{9}$$

This was noted by Gregg et al. (2006), who showed that we can estimate eq. (8) and then recover both the level and the percentage (multiplicative) effect. However, Gregg et al. (2006) defined the dif-in-dif "percentage method" as the percentage change in the treatment group minus the percentage change for the controls. This differs from $exp\left(\delta\right) - 1$. The reason is that the percentage change in the treatment group is equal to *%effect+%time+%effect×%time*. If we subtract the percentage change in the control group, we are left with *%effect×(1+%time)*. The difference is likely to be negligible if *%time* is small.

In spite of the equivalence in (9), we cannot interpret both $\tau$ and $\delta$ as causal effects. If we believe that the common trends assumption holds in multiplicative form, $\tau$ includes not only the level change due to the treatment, but also the difference between the time change in levels for the treatment and control groups. Indeed, $\tau$ is equal to the cross-difference from eq. (7), because the interaction term in the linear model identifies the difference in the change *in levels* of the two groups. Referring again to the example from Figure 1, this quantity is obtained by subtracting the incorrect linear counterfactual. The difference with respect to the true effect is in this case negative, because the correct counterfactual, calculated using a multiplicative trend, lies below the linear one.

More generally, the equivalence (9) does not work if we are willing to condition on other covariates, such as demographic controls. The reason is that the equation for the observed outcome is no longer saturated. Therefore it must be that either the linear model is correctly specified, or the exponential one, but not both. This is also true if we have more than two periods and a time trend is included.

The discussion of how the different specifications of time effects are crucial for

causal interpretation is related to Angrist and Pischke (2009, pg. 230) comment that the assumption of common trends can hold either in logs or in levels, but not in both. We find it more natural to look at the choice between multiplicative or additive effects, rather than focusing on whether taking logs or not. This perspective has the advantage of stressing the distinction between specification and estimation. More importantly, in the next section we show that the multiplicative model and the log-linearised one are equivalent only under a strong restriction.

## 2.2   Estimation of a multiplicative dif-in-dif model

A popular solution to estimate a multiplicative model is to log-linearise it. In this section we show that the standard mean-independence assumption imposed on the error tem of the original exponential model is not enough to guarantee the unbiasedness of the OLS estimator of the log-linearised version. The reason is that the mean-independence of the multiplicative error term in the exponential model does not imply the mean-independence of the (log-linearised) error in the log-linearised model. An alternative is to directly estimate the non-linear model.

To discuss this issue in details, we follow Santos Silva and Tenreyro (2006), but we reinterpret the problem in the context of the dif-in-dif setting. In line with the previous section, define a mean independent error term $\eta_{it}$ that enters multiplicatively in the model for the observed outcome (eq. 4), so that $y_{it} = exp\left(\cdot\right)\eta_{it}$ with $E\left[\eta_{it}|1, treated_{it}, post_{it}\right] = 1$. To estimate it, we can log-linearise:

$$lny_{it} = \beta_0 + \beta_1 treated_{it} + \beta_2 post_{it} + \delta treated_{it} \times post_{it} + ln\eta_{it}. \qquad (10)$$

In order for the OLS estimator on this log-linearised model to consistently estimate $\delta$, we need $E\left[ln\eta_{it}|1, treated_{it}, post_{it}\right] = 0$. However, as argued by Santos Silva and Tenreyro (2006) and Blackburn (2007), the mean independence assumption imposed on the original exponential model does not ensure that $ln\eta_{it}$ is mean-independent

as well. The reason is Jensen's inequality, which implies that the mean of the log is not equal to the log of the mean. In general, $E\left[ln\eta_{it}|1, treated_{it}, post_{it}\right] = 0$ would hold if we could impose the stronger condition that $\eta_{it}$ was statistically independent from $x_{it} \equiv (1, treated_{it}, post_{it})$. This condition would imply:

$$Var\left[y_{it}|1, treated_{it}, post_{it}\right] =$$

$$\sigma_{it}^2 exp\left(2\beta_0 + 2\beta_1 treated_{it} + 2\beta_2 post_{it} + 2\delta treated_{it} \times post_{it}\right) \quad (11)$$

where $\sigma_{it}^2 = Var\left(\eta_{it}\right)$. In order for the error term of the log-linearised model to be mean-independent, the ratio of variances between different groups or time periods should be directly related to the differences in the conditional mean. Furthermore, the treatment effect must not only shift the conditional mean, but also increase (or decrease) the conditional variance by a factor equal to the square of $exp\left(\delta\right)$. This pattern of variance does not necessarily hold under the weaker condition of mean independence ($E\left[\eta_{it}|x_{it}\right] = 1$), from which we started.

For instance, take a multiplicative error $\eta_{it}$ such that $E\left[\eta_{it}|x_{it}\right] = 1$, but suppose that the stronger statistical independence condition $\eta_{it} \perp\!\!\!\perp x_{it}$ holds only in the absence of the treatment, that is when $treated_{it} \times post_{it} \neq 1$. Differently, assume that the treatment has a distributional effect which differs from the simple increase in variance that would follow from condition (11), so that the error $\eta_{it}$ is not statistically independent. Following the previous discussion, this causes the log-linearised error $ln\eta_{it}$ to be not mean-independent, because its conditional mean now depends on $treated_{it} \times post_{it}$. The conditional expectation of $lny_{it}$ becomes:

$$E\left[lny_{it}|1, treated_{it}, post_{it}\right] =$$

$$\beta_0^* + \beta_1 treated_{it} + \beta_2 post_{it} + \delta^* treated_{it} \times post_{it} \quad (12)$$

$$\beta_0^* = \beta_0 + E\left[ln\eta_{it}|treated_{it} \times post_{it} \neq 1\right] \tag{13}$$

$$\delta^* = \delta + E\left[ln\eta_{it}|treated_{it} \times post_{it} = 1\right] - E\left[ln\eta_{it}|treated_{it} \times post_{it} \neq 1\right]. \tag{14}$$

The OLS estimator for the log-linearised model would therefore be consistent for $\delta^*$, which differs from the true $\delta$ because it confounds distributional with mean effects. The bias is equal to the difference between the mean of $ln\eta_{it}$ in the treated group in the post period and the one for the control group and pre-policy period.

It should be added that this bias would be present even if the true treatment effect on the mean was zero, and there was no difference across groups or time ($\beta_1 = \beta_2 = 0$). A similar bias would arise if the treatment had no effect at all on the outcome distribution, but in the second period there was some change in the variance of $y$ within the treatment group that violates the assumption $\eta_{it} \perp\!\!\!\perp x_{it}$. Such a situation would be compatible with the multiplicative common trends assumption stated in terms of conditional mean (eq. 2), because it does not impose any restriction on higher moments.[7]

Differently, the OLS estimator might not be affected by a situation as in Blackburn (2007), where the conditional variance across groups does not follow the pattern in eq. (11), but the condition is respected over time within the same group. Suppose that $E\left[\eta_{it}|x_{it}\right] = 1$. However, assume that the variance and higher moments in the distribution of $\eta_{it}$ depend on the group, though neither on the time period, nor on the treatment. In general, we would have that

$$E\left[ln\eta_{it}|treated_{it} = 0, post_{it} = 0\right] = E\left[ln\eta_{it}|treated_{it} = 0, post_{it} = 1\right]$$

$$\neq E\left[ln\eta_{it}|treated_{it} = 1, post_{it} = 0\right] = E\left[ln\eta_{it}|treated_{it} = 1, post_{it} = 1\right]. \tag{15}$$

---

[7]As discussed by Abadie (2005), if the two groups are unbalanced in terms of observable covariates, this may generate differential time trends. Interestingly, these non-parallel dynamics may involve not only the average, but also higher moments. In this case, the method suggest by Abadie (2005) applied to the log-linearised model may also reduce the bias introduced by the differential pattern in the variance. However, it should be noted that our discussion holds even if covariates are perfectly balanced, because the heteroskedasticity may be induced by the treatment itself or by other distributional changes across time that differ across groups.

Therefore, both the intercept and the coefficient on the group dummy will be different from $\beta_0$ and $\beta_1$, but the coefficient on the interaction would recover the true treatment effect.

Nevertheless, we know from the literature that there is an alternative estimation strategy which is consistent in both cases, because it only requires $\eta_{it}$ to be mean independent from $x_{it}$, and not necessarily statistically independent. Santos Silva and Tenreyro (2006) and Blackburn (2007) proposed to directly estimate the non-linear model. In practice, one can use both Non Linear Least Squares (NLS) and Poisson Pseudo Maximum Likelihood (PPML). In both cases, the estimator simply exploits the restriction that the conditional expectation is correctly specified as exponential. Santos Silva and Tenreyro (2006) argued in favour of the latter, because NLS is likely to be less efficient.

The PPML simply estimates the model as if it was a Poisson, by maximizing the relative likelihood.[8] The fact that the actual variable is continuous rather than count does not hinder the consistency of the estimator. This follows from the well known result that the Poisson Maximum Likelihood estimator is consistent as long as the mean is correctly specified, which is given by the standard mean independence assumption $E\left[\eta_{it}|1, treated_{it}, post_{it}\right] = 1$, or equivalently by eq. (4). However, given that the actual variable does not respect the other properties of the Poisson distribution, standard maximum likelihood inference would be invalid, and therefore a robust covariance matrix should be used.

Practically, PPML can be implemented in the most popular statistical packages and results can be easily interpreted. In Stata$^{\text{TM}}$, one can simply run the *poisson* command, with all variables in levels.

---

[8] An alternative and equivalent way to fit the model is by GLM with a log link function.

## 2.3 Heterogeneous treatment effects in the multiplicative case

One important question is what the level and multiplicative model are effectively identifying when treatment effects are heterogeneous, in the sense that they vary at the individual level.[9] In the level model, given the additive nature of the effects, the well know result is that the dif-in-dif estimand identifies the average treatment effect on the treated.

To discuss the multiplicative model, we follow the IV-exponential model in Angrist (2001) and we include an individual fixed effect $\omega_i$ and a heterogeneous treatment effect $\delta_i$. Here the superscript $^*$ is used to differentiate the error $\eta_{it}^*$ in the unobservable model from the error $\eta_{it}$ in the observable model:

$$y_{1igt} = y_{0igt} exp\left(\delta_i\right) = exp\left(\lambda_t + \omega_i + \delta_i\right)\eta_{it}^*, \tag{16}$$

$$E\left[\eta_{it}^*|t, \omega_i, \delta_i\right] = 1. \tag{17}$$

Similarly to the additive dif-in-dif model, the individual fixed effects $\omega_i$ are introduced to allow a potential source of heterogeneity between the two groups. Differently, the purpose of $\delta_i$ is to allow for the effect of the treatment to differ across individuals and groups.[10]

Regarding the individual fixed effects, there is nothing new with respect to the discussion from section 2.1. In the cross-sectional case, in order for the dif-in-dif estimator to remove the heterogeneity in the $\omega_i$ we need the composition of both groups not to change over time, so that the expected values of the individual fixed effects are stable (Blundell and Macurdy, 1998). However, it is interesting to understand whether the estimator identifies the average of the treatment effects $\delta_i$

---

[9]We do not focus on the heterogeneity with respect to different observable characteristics. This discussion would hold even if the treated and control groups were completely homogeneous in terms of observables, but the distribution of the treatment effect was heterogeneous across individuals.

[10]Blackburn (2007, pg. 91-92) discussed how to include individual fixed effects in an exponential conditional mean, in order to correctly estimate wage differentials. Differently from his paper, we also allow individual heterogeneity in the treatment effects and we mainly focus on the estimation of their average.

or something different.

The (unobservable) model generating the observable exponential model from eq. (4) becomes

$$y_{it} = exp\left(\lambda_t + \omega_i + \delta_i \mathbf{1}\left[g = treated, t = post\right]\right)\eta_{it}^*. \tag{18}$$

If we use PPML to estimate it, we know that $exp\left(\delta\right)$, the exponentiated coefficient on the interaction term, identifies the ROR (eq. 6). Given the model in (18), it follows that

$$exp\left(\delta\right) - 1 =$$
$$\frac{exp\left(\lambda_{post}\right)E\left[exp\left(\omega_i\right)exp\left(\delta_i\right)|g = treated\right]}{exp\left(\lambda_{pre}\right)E\left[exp\left(\omega_i\right)|g = treated\right]} \Big/ \frac{exp\left(\lambda_{post}\right)E\left[exp\left(\omega_i\right)|g = control\right]}{exp\left(\lambda_{pre}\right)E\left[exp\left(\omega_i\right)|g = control\right]} - 1$$
$$= \frac{E\left[y_{1igt}|g = treated\right] - E\left[y_{0igt}|g = treated\right]}{E\left[y_{0igt}|g = treated\right]} \tag{19}$$

so that only the multiplicative treatment effect on the average is identified, and not the average of the multiplicative effect. This was already noted by Angrist (2001) for IV estimation of an exponential model.

If we assume the stronger condition that $\eta_{it}^* \perp\!\!\!\perp (t, \omega_i, \delta_i)$, we can log-linearise the model in (18):

$$lny_{it} = \lambda_t + \omega_i + \delta_i \mathbf{1}\left[g = treated, t = post\right] + ln\eta_{it}^* \tag{20}$$

where statistical independence ensures that $ln\eta_{it}^*$ is mean-independent from $t$, $\omega_i$ and $\delta_i$, so that (20) represents a conditional expectation. In this case, the standard dif-in-dif level regression applied to $lny_{it}$ identifies the cross difference (in levels) of

the outcome (eq. 7), which from model (20) is equal to

$$\{\lambda_{post} + E\left[\omega_i | g = treated\right] + E\left[\delta_i | g = treated\right] - \lambda_{pre} - E\left[\omega_i | g = treated\right]\}$$

$$- \{\lambda_{post} + E\left[\omega_i | g = control\right] - \lambda_{pre} - E\left[\omega_i | g = control\right]\} =$$

$$= E\left[\delta_i | g = treated\right]. \quad (21)$$

This quantity, although related to the original parameters, is not of direct interest. The reason is that the multiplicative treatment effect for each individual is equal to $exp(\delta_i) - 1$, but in general $E\left[\delta_i | g = treated\right] \neq E\left[exp(\delta_i) - 1 | g = treated\right]$. The problem is that statistical independence of the error term holds in the *unobserved* model (18). Differently, the error term $\eta_{it}$ in the *observed* model is likely to be heteroskedastic, as it depends on the distribution of the individual treatment effects $\delta_i$ within each group. Using log-linearisation we can recover the average causal effect on the logs $E\left[\delta_i | g = treated\right] = E\left[lny_{1igt} - lny_{0igt} | g = treated\right]$ but we cannot use it to go back to the multiplicative effect on the average in the original scale (19).

In a nutshell, in the presence of heterogeneous effects the standard mean independence assumptions behind the exponential dif-in-dif model only allow us to identify the multiplicative effect on the average for the treated group, and not an average multiplicative effect. The presence of heterogeneous treatment effects is likely to induce a dependence between the error term $\eta_{it}$ in the observed model and the covariates. Therefore the statistical independence assumption is not likely to hold, and the OLS estimates of the log-linearised model would not recover the quantity of interest.

# 3   An applied example

To provide an example, we apply the PPML estimator in a dif-in-dif setting to assess the effects of the recent introduction of the Educational Maintenance Allowance

(EMA) in the United Kingdom on household expenditures. EMA provided up to £30 per week to 16-18 year olds in low income households, conditional on them attending a full time educational course. The policy was introduced nationwide in September 2004, although a pilot took place in 1999-2000. Dearden et al. (2009) provide evidence suggesting that EMA pilots increased post 16 schooling by 5-7 percentage points. Further details of the reform can be found in Appendix B.

Here we study how families targeted by the scheme spent the available resources. In line with the theory of the previous section, we specify a multiplicative model for expenditure, using both OLS on the log-linearised values and PPML.

## 3.1   Data and identification strategy

We take advantage of expenditure data from the first five years of the Expenditure and Food Survey (EFS). Details of the survey are available in Appendix B. All income and expenditure figures are in weekly equivalents and are expressed in December 2005 terms using the retail price index, available from the Office for National Statistics.

The estimation sample consists of all households with at least one child aged either 14, 15 or 17. The EFS operates on the basis of a financial year (April – March) and our sample includes periods from April 2001 until March 2006. The reform coincides with the start of the school year in 2004 (September). Given that we have more than two time periods, we include a full set of year dummies, plus a set for the month of interview to account for seasonality. The departure from the simple $2 \times 2$ setting implies that the exponential model is not simply a reparametrisation of the level one.

Treatment and control groups are not defined according to information on education status, which may be endogenous to the reform. Rather, information is used on exogenous age at interview of the household members. The treated group of households is defined to be those where at least one 17 year old is residing, be-

cause conditional on having low income they will be eligible to receive EMA. The control group is formed of households where at least one 14 or 15 year old resides, excluding households defined to be in the treated group as above. We exclude 16 and 18 year olds to avoid misclassification, because full information on date of birth would be required to determine EMA eligibility status and the EFS only contains information on age at interview. Table B.2 in appendix B demonstrates that, under this definition, treatment and control groups are similar in terms of observable characteristics.

In line with the spirit of this paper, we focus on the issues related to functional form specification and estimation of a multiplicative model. Clearly, accounting properly for the time trend does not imply that estimates can be interpreted as causal effects. Other problems, such as the presence of an Ashenfelter's dip or of anticipation effects, can bias the results and are discussed in details in Appendix B. We report the robust Huber-White standard errors.[11]

EMA was already in operation in 41 of the 150 pilot Local Education Authorities before the start of our sample period. Pilot areas cannot be removed from the treated group as the EFS does not record information on LEA status. This implies that treatment is less than 100% for the treated group and that the presented estimates of the effect of the policy on household expenditure patterns, therefore, represent a lower bound of the effect on those actually receiving EMA. On the one hand, it may be enough for a policy maker to know this intent to treat, on the other, the interest may be in the effect for those that actually received EMA. In Appendix B we comment on the possibility of rescaling the estimated treatment effects to reflect this fact. We show that in the case of a multiplicative model, rescaling may be problematic.

---

[11]Procedures to correct for the fact that regular standard errors may overstate the precision of estimates of a treatment effect in dif-in-dif regressions are the subject of an ongoing debate (see Bertrand et al., 2004; Donald and Lang, 2007; Wooldridge, 2003, 2006). To address problems of serial correlation, we restrict the sample to only 5 years of data. Furthermore, under the definition of the treatment and control groups, we see no reason to believe that there are shocks that occur at the group level.

We present estimates for 7 major areas of spending: food and non-alcoholic drinks; alcoholic beverages and tobacco; clothing and footwear; furnishings, household equipment and carpets; transport; communication; and recreation. Following on from the earlier discussion, it is natural to specify the common trends assumption in multiplicative form. That is expenditures, following the growth of the economy, increase by a constant percentage in the absence of treatment.

In general, we know that total household expenditure tends to be log-normally distributed (see, for instance, Battistin et al., 2009). One might claim that, consequently, log-linearisation is harmless. This is not necessarily the case. First of all, the required assumption refers to the conditional distribution of $y_{it}$, that is within group and time period. Secondly, even if the error $\eta_{it}$ was log-normal, we would need it to be statistically independent from the covariates. Indeed, in the simulation in section 4 we show that a log-normally distributed disturbance is not enough for log-OLS consistency. Finally, nothing ensures log-normality of each category of spending. This is particularly true if there is a non-negligible proportion of zero, which may be due to measurement error in the recording of small amounts, as pointed out by Santos Silva and Tenreyro (2006) for trade data.

## 3.2   Results

Table 1 presents dif-in-dif estimates of the effect of the EMA scheme on each of the 7 spending categories for the treated group of households. The results in columns 1 correspond to estimates of the multiplicative effect using OLS on the log-linearised model, while in column 2 the reform effect is estimated directly using the PPML estimator. For completeness, we also report OLS estimates for a level model of expenditure in column 3. It is important to stress that, as usual, observations with zero expenditure are dropped from OLS log estimates, while not for the PPML estimates. Nevertheless, results are similar when excluding zero cases across all estimators or when setting the logarithm equal to zero in the case of a zero expenditure (results

available on request).

Following the national roll out of EMA in September 2004, we expect the treated group of households to increase expenditures in some categories. For the OLS log-linearised estimates in column 1 we see positive dif-in-dif estimates for food, non-alcoholic drinks; alcoholic beverages and tobacco; transport; recreation and negative effects for clothing footwear; furnishings household equipment and carpets; communication. None of the estimated effects are, however, statistically significant.

Turning to the reform effects in column 2, EMA might also have distributional effects that make the multiplicative error term statistically dependent on the time and group dummies. In this case, we expect the previous OLS log results to suffer from bias, while PPML results should be consistent. For most of the categories, coefficients are in line with the OLS log results, but for transport spending the estimated coefficient has increased in magnitude. Moreover, it is now statistically different from zero at the 5% level. The result implies an increase around 23% in transport spending due to the reform, calculated as $exp\left(\widehat{\delta}\right) - 1$. This finding is in line with evidence from the EMA piloting, in which EMA recipients were more likely to be contributing to transport expenditures compared to non-recipients and EMA eligibles residing in control areas (see Ashworth et al., 2001, p. 59). In comparison to the standard log expenditure estimates of column 1, the PPML coefficient implies an EMA effect of 10.8 percentage points bigger, which is more precisely estimated. For the remaining spending categories, we observe statistically insignificant coefficients, which are also generally smaller than the effect on transport.[12]

On the OLS level results of column 3, the estimated signs and significance of the interaction terms match with the PPML results. For transport spending, the *treated* × *post* interaction is a statistically significant £16.46, which corresponds to

---

[12]One criticism might be that the result for transport is accidental, because in a full set of regressions it is not unlikely to find at least one statistically significant estimate. However, here we focus on the difference between OLS (logs) and PPML results. Moreover, given the relatively small sample sizes, rather than making adjustments to standard errors, which can be conservative and computationally intensive (Duflo et al., 2008), we draw on the evidence from the EMA trials to support our conclusions.

Table 1: Estimates of the EMA Effect on 6 Major Expenditure Categories

| | Full Sample | | | Low Income Sample | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS Log | PPML | OLS Level | OLS Log | PPML | OLS Level |
| **Food and** | 0.069 | 0.026 | 1.777 | 0.039 | 0.010 | 0.584 |
| **Non-alcoholic drinks** | (0.051) | (0.041) | (2.838) | (0.080) | (0.065) | (3.868) |
| Observations | 2626 | 2631 | 2631 | 1314 | 1317 | 1317 |
| Reset(p-value) | 0.0927 | 0.7315 | 0.8800 | 0.5150 | 0.5410 | 0.5940 |
| **Alcoholic Beverages** | 0.026 | 0.104 | 1.276 | -0.022 | 0.178 | 2.458 |
| **and Tobacco** | (0.104) | (0.101) | (1.659) | (0.156) | (0.152) | (2.320) |
| Observations | 1968 | 2631 | 2631 | 935 | 1317 | 1317 |
| Reset(p-value) | 0.8720 | 0.2829 | 0.1083 | 0.4835 | 0.0107 | 0.0417 |
| **Clothing and** | -0.137 | 0.012 | -0.477 | -0.136 | -0.024 | -1.638 |
| **Footwear** | (0.100) | (0.091) | (4.060) | (0.152) | (0.142) | (5.350) |
| Observations | 2342 | 2631 | 2631 | 1128 | 1317 | 1317 |
| Reset(p-value) | 0.8158 | 0.9232 | 0.3009 | 0.0337 | 0.8239 | 0.3901 |
| **Furnishings, HH** | -0.126 | -0.092 | -3.511 | -0.150 | -0.211 | -6.015 |
| **Equipment, Carpets** | (0.115) | (0.177) | (7.108) | (0.170) | (0.219) | (6.976) |
| Observations | 2575 | 2631 | 2631 | 1282 | 1317 | 1317 |
| Reset(p-value) | 0.2364 | 0.5347 | 0.3370 | 0.8011 | 0.1806 | 0.1282 |
| **Transport** | 0.116 | 0.208** | 16.458** | 0.127 | 0.328** | 18.292* |
| | (0.098) | (0.095) | (8.276) | (0.153) | (0.163) | (9.851) |
| Observations | 2505 | 2631 | 2631 | 1204 | 1317 | 1317 |
| Reset(p-value) | 0.7219 | 0.9051 | 0.6783 | 0.0936 | 0.7100 | 0.3995 |
| **Communication** | -0.083 | -0.032 | -0.553 | -0.011 | 0.029 | 0.559 |
| | (0.064) | (0.074) | (1.406) | (0.095) | (0.118) | (2.027) |
| Observations | 2548 | 2631 | 2631 | 1248 | 1317 | 1317 |
| Reset(p-value) | 0.4642 | 0.3287 | 0.2247 | 0.6068 | 0.1991 | 0.1732 |
| **Recreation** | 0.014 | -0.010 | -0.633 | 0.083 | 0.204 | 12.676 |
| | (0.081) | (0.092) | (7.745) | (0.118) | (0.150) | (9.522) |
| Observations | 2629 | 2631 | 2631 | 1315 | 1317 | 1317 |
| Reset(p-value) | 0.8858 | 0.1458 | 0.0923 | 0.4734 | 0.0194 | 0.0131 |

Note: * p<.10, ** p<.05, *** p<.01. Standard errors (robust) in parentheses.

18.3% of the pre-reform mean. However, the dif-in-dif coefficient presented only corresponds to the causal effect of EMA on the level of expenditure if we are willing to impose common trends in expenditure levels. If the common multiplicative trend is the correct one, then no meaningful interpretation can be given to the coefficient of the level model.

Columns (4)-(6) better target the groups affected by the reform by repeating the analysis for a sample of low income households. Results give further strength to the main finding with PPML estimates in column 5 suggesting that households devoted the additional resources from EMA primarily to transport spending. The PPML estimate increases in magnitude with little reduction in the precision (comparing to column 2). This is once again in contrast to the OLS log result which remains smaller and statistically insignificant.

We compare how the models perform on Ramsey's RESET test (Ramsey, 1969) for misspecification of the conditional mean. It involves calculating the square of the fitted values and including them as an additional regressor. P-values for the significance of this coefficient are reported alongside the main results in table 1. For the full sample, no evidence is found of misspecification of the conditional mean. For the low income sample the same picture emerges; however, the test is only marginally passed in the case of the PPML and level estimates for alcoholic beverages and tobacco; and recreation; and the OLS log results for clothing and footwear.

The results of the test illustrate an important point that correct specification is not sufficient for causal interpretation. For example, in the simple 2-period-2-group case, the conditional expectations of both $lny$ and $y$ are correctly specified as linear because the model is saturated. Our choice about which estimates to interpret as causal effects critically depends on our belief about the nature of common trends. Furthermore, under heteroskedasticity the effect estimated with logs might confound mean with distributional effects, even though the model for $lny$ is correctly specified.

# 4 A simulation

In order to illustrate the theoretical arguments, this section reports results from monte-carlo simulations. A group of hypothetical reforms are considered that have common multiplicative time and treatment effects on the mean, but that differ in terms of their distributional impacts. We first focus on the difference between the PPML on the original exponential model and the OLS estimator on the log linearised one. Results show that the former is robust to a change in variance that introduces heteroskedasticity across groups over time, while the latter is biased. We then consider what happens if we estimate a standard additive dif-in-dif regression, ignoring the fact that the true model is multiplicative.

In the simulation we consider a setting similar to that reported in figure 1, so that the reader can follow the difference between the estimates in levels and the ones in multiplicative form (see Appendix C for a graph using the simulation parameters). The outcome of interest is generated according to equation (4), i=1,..., 2631. The sample size and size of the groups are selected to be as in the applied example of section 3, in order to be able to detect similar differences between estimators.

Each replication is generated according to $\beta_0 = 3.5$, $\beta_1 = -0.4$, $\beta_2 = 0.03$ and with the hypothetical reform having a constant multiplicative treatment effect (ratio of ratios) equal to $\delta = 0.2$. Simulations with a negative time trend and a positive difference between groups lead to the same conclusions and are available on request.

A mean-independent random error term is introduced, so that each individual observation is generated according to $y_{it} = exp(x_{it}\beta)\eta_{it}$ where $\eta_{it}$ is a log normal random variable with $E(\eta_{it}|x) = 1$ and $var(\eta_{it}|x) = \sigma_{it}^2$. The variance $\sigma_{it}^2$ is specified as:

$$\sigma_{it}^2 = exp(\alpha \times \mathbf{1}(treated_{it} \times post_{it} = 1)) \tag{22}$$

where $\mathbf{1}$ is an indicator function and $\alpha$ a parameter that determines the degree

27

of heteroskedasticity in $\eta_{it}$.

To assess the performance of the three estimation strategies outlined above, simulations are reported for five key values of $\alpha$. Table 2 reports results from 1000 replications of the simulation procedure.

The first special case of interest is the scenario where $\alpha = 0$ implying $\eta_{it}$ is statistically independent of the treatment and other regressors, so that OLS estimates from the log-linear model will provide consistent estimates of the multiplicative treatment effect. As expected both the OLS log in column 1 of Table 2 and PPML estimates in column 2 are close to the true multiplicative treatment effect of 0.2. Whilst the difference between the two estimates is negligible, the OLS log estimates are less dispersed, confirming the greater efficiency of the OLS estimator under statistical independence of the error term.

With $\alpha = 0.1$ we introduce heteroskedasticity. For instance, this could be the case if the treatment has a distributional effect above that due to the simple increase in the conditional mean. This may also be the consequences of other changes in the higher moments of the distribution over time. Indeed, it is important to remember that the standard dif-in-dif identifying assumption of common trends (in this case multiplicative) expressed in terms of the conditional mean of $y_{it}$ places no restriction on these moments. From eq. (14) we know that the bias in using OLS on the log-linearised outcome is due to the fact that the differential increase in variance causes the log-linearised error to be not mean independent. In this case, the analytical bias depends on the formula for the variance of $\eta_{it}$ (eq. 22) and is equal to[13]

$$E\left[ln\eta_{it}|treated_{it} \times post_{it} = 1\right] - E\left[ln\eta_{it}|treated_{it} \times post_{it} \neq 1\right] =$$
$$- 0.5\left[ln\left(1 + exp(\alpha)\right)\right] + 0.5\left[ln\left(2\right)\right]. \quad (23)$$

Therefore, we expect that an increase in variance in the post-treatment period for

---

[13]The formula is obtained by the equations relating the moments of a log-linear random variable $X$ with those of $lnX$.

Table 2: Simulation results

|  | (1) OLS Log | (2) PPML | (3) OLS Level |
|---|---|---|---|
| $\alpha = 0$ | | | |
| Treated x Post | .1999924 | .1984138 | 4.689477 |
|  | (0.0711943) | (0.0842905) | (2.510064) |
| Treated | -.399073 | -.3976382 | -10.84626 |
|  | (0.0398905) | (0.0474284) | (1.290385) |
| Park Test | 0.052 | 0.014 | - |
| Breusch-Pagan | 0.049 | - | - |
| $\alpha = 0.1$ | | | |
| Treated x Post | .1710568 | .1951523 | 4.585499 |
|  | (0.0731679) | (0.0882464) | (2.62745) |
| Treated | -.4002561 | -.4000494 | -10.90821 |
|  | (0.0404467) | (0.0470304) | (1.277552) |
| Park Test | 0.04 | 0.005 | - |
| Breusch-Pagan | 0.154 | - | - |
| $\alpha = 0.2$ | | | |
| Treated x Post | .1506154 | .2026701 | 4.814591 |
|  | (0.0727646) | (0.0887447) | (2.625155) |
| Treated | -.4010721 | -.401937 | -10.96927 |
|  | (0.0406465) | (0.0496038) | (1.345574) |
| Park Test | 0.046 | 0.017 | - |
| Breusch-Pagan | 0.441 | - | - |
| $\alpha = 0.4$ | | | |
| Treated x Post | .0861068 | .1960358 | 4.632431 |
|  | (0.0776181) | (0.096877) | (2.870258) |
| Treated | -.399795 | -.3997392 | -10.90645 |
|  | (0.0404998) | (0.0484243) | (1.310166) |
| Park Test | 0.064 | 0.022 | - |
| Breusch-Pagan | 0.945 | - | - |
| Note: Results from 10,000 replications of the simulation procedure described in section 4. Mean of the estimated coefficients reported with standard deviations in parentheses. | | | |

the treated group (due to $\alpha > 0$) should induce a negative bias. Accordingly, in Table 2 we observe that now the OLS log procedure performs less well. Here, the OLS log estimates confound the distributional effect of treatment with the mean effect and are biased for the true multiplicative effect. The distance from the true effect is around 2.9 percentage points, similar to the 2.6 point difference that can be calculated using the formula from (23).[14] As expected, the value of this bias increases with the value of $\alpha$, even though the variance of the estimated effects remains small. For example, the mean of the estimated treatment effects being only 43% of the true effect when $\alpha = 0.4$. On the other hand, the PPML estimator performs well under all values of $\alpha$, giving estimates close to the true treatment effect in all cases.

It is worth pointing out that the parameter values considered above imply an independent effect of treatment on the conditional variance of $y$ that deviates only slightly from statistical independence. For example, under the strongest pattern of heteroskedasticity considered ($\alpha = 0.4$), the independent effect of treatment is to increase the conditional standard deviation of $y$ by only 22%, whereas when $\alpha = 0.1$ the increase in standard deviation is just 5%. Even when these very small distributional effects of treatment are introduced, the estimates in the table from the log-linearised model are strongly biased.

Another important question is what would happen if, ignoring the multiplicative structure, we estimate a standard additive dif-in-dif regression. The set of parameters imply that, for the treated group, the difference in levels between $y_{1i}$ and $y_{0i}$ (the counterfactual) in the post treatment period would be £5.06, as could be easily calculated from the exponential model. This is the part of the change that can be attributed to the reform after accounting properly for the multiplicative trend. Differently, the estimand of a standard dif-in-dif in levels would be $[exp(3.5 - 0.4 + 0.03 + 0.2) - exp(3.5 - 0.4)] - [exp(3.5 + 0.03) - exp(3.5)]$, which

---

[14]The value of the bias is close to the theoretical one also in the simulations for larger values of $\alpha$.

is £4.73. The difference is due to the fact that the multiplicative time trend is not properly accounted for. However, given that the change over time is relatively small, the distortion is not large.

The simulation result illustrates this point. Column 3 of table 2 presents estimates from standard OLS estimation in levels. From the table, we observe that the estimated treatment effects repeatedly underestimate the true reform impact. The effect is £4.69 in the baseline case, in contrast to the change in levels that is implied by the multiplicative model (£5.06). The bias is independent from the value of $\alpha$. So although the regression for $y_i$ is saturated and therefore correctly specified, the estimated effect in levels confounds the treatment and trend effects.

Given an exponential model for the conditional mean of $y$, a researcher may wish to test whether estimation by log-linearisation will be consistent. Table 2 also presents evidence on the performance of two tests. The first is a Park test (Manning and Mullahy, 2001; Santos Silva and Tenreyro, 2006) for whether the conditional variance of $y$ is proportional to the conditional mean squared. This involves testing whether $\gamma_1$ is statistically different from 2 in the equation:

$$(y_i - \widetilde{y}_i)^2 = \gamma_0(\widetilde{y}_i)^{\gamma_1} + \epsilon_i \tag{24}$$

where $\widetilde{y}$ is a consistent estimate of $E[y_i|x]$, obtained using PPML. Inference from equation (24) uses the Eicker-White robust covariance matrix estimator. The second test is a standard Breusch-Pagan (BP) test for whether the estimated variance of the residuals from the log model is statistically dependent on the value of the *treated* × *post* variable. It is the optimal test for heteroskedasticity when errors are normally distributed. For both tests, rejection rates at the 5% level are reported in the table.

Equation (24) is estimated by both log-linearisation (column 1) and directly by PPML (column 2). Results from the simulation are not promising. In all cases the Park test fails to detect the mild pattern of heteroskedasticity that treatment intro-

31

duces into the model. The rejection rates are around 5% for all values of $\alpha$ (When stronger patterns of heteroskedasticity were introduced to the model, for example with a simulation with a constant variance of $y$, the performance of the Park test improves with rejection rates reaching 72%). For the BP test, the heteroskedasticity introduced into $\eta_i$ is detected with reasonable power. For example, in the case where $\alpha=0.4$ the test detects the inadequacy of the log-linearised specification 94.5% of the time. These results suggest that testing for heteroskedasticity in the model for $lny$ with a BP test can be informative when deciding upon an estimation strategy with a multiplicative model in the dif-in-dif setting.

In appendix C we also analyse the case in which the treatment has no distributional effect, but the pattern of variance across the treated and control groups does not respect the proportional structure from eq. (11). As discussed in the theoretical section, in this case the log-OLS estimator for the treatment effect is consistent, while the treated-control difference is biased. We also analysed the case with a constant variance of $y$. Again here log-OLS for the treatment effect performs poorly. Results are available from the authors.

# 5   Conclusion

We critically assessed the standard practice of log-linearising in a dif-in-dif setting. We argued that a researcher should first decide whether a multiplicative or additive effect model is appropriate for the non-transformed outcome, because we cannot give a causal interpretation to both. If the multiplicative model is chosen and the researcher makes only a standard mean independence assumption, using Poisson Pseudo Maximum Likelihood on the non-transformed variable can be preferable to using OLS on the log-linearisation. The reason is that the latter might give biased estimates of the multiplicative effect if there are changes in the higher moments of the outcome distribution that make the log-linearised error not mean independent.

In particular, this bias may cause the OLS estimator to confound other distributional effects with the treatment effect on the mean.

As a summary, we think that the best practice for an applied researcher willing to estimate a dif-in-dif model with continuous outcome should be (a summary table can be found at the end of Appendix C):

1. Decide whether the time trend is more likely to hold in multiplicative or in level form.

2. If in levels, the best solution would be to use the standard level model and estimate it through OLS. The coefficient on the interaction term could be interpreted as an average treatment effect for the treated.

3. If in multiplicative form, the most coherent solution is to use and estimate an exponential model, with a multiplicative treatment effect.

   (a) Without covariates, the multiplicative treatment effect can be recovered from OLS estimates of the standard dif-in-dif regression in levels (eq. 9).

   (b) Estimating the exponential model with PPML allows for covariates and for the presence of zeros in the dependent variable, and does not require statistical independence of the error term.

   (c) The researcher can test for heteroskedasticity using a BP test for the presence of heteroskedasticity with respect to the $treated \times post$ variable. If they fail to reject the null of homoskedasticity, and the researcher is willing to assume statistical independence, OLS on the log-linearised model would be unbiased and efficient. This method also requires to eliminate or censor the zeros, which may introduce another source of bias.

   (d) In the case of heterogeneous effects, the exponentiated coefficient on the interaction term $(exp\,(\delta)-1)$, can be interpreted as a multiplicative effect on the average for the treated group.

# References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *Review of Economic Studies 72*, 1–19.

Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters 80*, 123–129.

Ai, C. and E. C. Norton (2008). A semiparametric derivative estimator in log transformation models. *Econometrics Journal 11*, 538–553.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business and Economic Statistics 19*, 2–16.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics*. Princeton: Princeton University Press.

Ashworth, K., J. Hardman, W. Liu, S. Maguire, S. Middleton, L. Dearden, C. Emmerson, C. Frayne, A. Goodman, H. Ichimura, and C. Meghir (2001). Education Maintenance Allowance: The First Year - A Quantitative Evaluation. Department for Education and Employment (DfEE) Research Report No. RR257.

Athey, S. and G. W. Imbens (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica 74*, 431–497.

Battistin, E., R. Blundell, and A. Lewbel (2009). Why is consumption more log normal than income? Gibrat's law revisited. *Journal of Political Economy 117*, 1140–1154.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*, 249–275.

Blackburn, M. L. (2007). Estimating wage differentials without logarithms. *Labour Economics 14*, 73–98.

Blundell, R. and T. E. Macurdy (1998). Labor supply: a review of alternative approaches. IFS Working Paper Series No. W98/18.

Buis, M. L. (2010). Stata tip 87: Interpretation of interactions in non-linear models. *The Stata Journal 10*, 305–308.

Dearden, L., C. Emmerson, and C. Meghir (2009). Conditional cash transfers and school dropout rates. *Journal of Human Resources 44*.

DellaVigna, S. and E. Kaplan (2007). The fox news effect: Media bias and voting. *The Quarterly Journal of Economics 122*, 1187–1234.

DiNardo, J. and D. S. Lee (2011). Program evaluation and research designs. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, pp. 463–536. Oxford: Elsevier.

Donald, S. G. and K. Lang (2007). Inference with Difference-in-Differences and Other Panel Data. *The Review of Economics and Statistics 89*, 221–233.

Duflo, E., R. Glennerster, and M. Kremer (2008). Using Randomization in Development Economics Research: A Toolkit. In T. P. Schultz and J. A. Strauss (Eds.), *Handbook of Development Economics*, pp. 3895–3962. Oxford: Elsevier.

Finkelstein, A. (2007). The aggregate effects of health insurance: Evidence from the introduction of medicare. *The Quarterly Journal of Economics 122*, 1–37.

Gregg, P., J. Waldfogel, and E. Washbrook (2006). Family expenditures post-welfare reform in the UK: Are low-income families starting to catch up? *Labour Economics 13*, 721–746.

Karaca-Mandic, P., E. C. Norton, and B. Dowd (2012). Interaction Terms in Non-linear Models. *Health Services Research 47*, 255–274.

Manning, W. G. (1998). The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of Health Economics 17*, 2983–295.

Manning, W. G. and J. Mullahy (2001). Estimating log models: to transform or not to transform? *Journal of Health Economics 20*, 461–494.

Mullahy, J. (1997). Instrumental-variable estimation of count data models: Applications to models of cigarette smoking behavior. *The Review of Economics and Statistics 79*, 586–593.

Mullahy, J. (1999). Interaction effects and difference-in-difference estimation in log-linear models. NBER Technical Working Paper 245.

Nunn, N. and N. Qian (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *The Quarterly Journal of Economics 126*, 593–650.

Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters 115*, 85–87.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological) 31*, 350–371.

Santos Silva, J. M. C. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*, 641–658.

Wooldridge, J. M. (2003). Cluster-Sample Methods in Applied Econometrics. *American Economic Review 93*, 133–138.

Wooldridge, J. M. (2006). Cluster-Sample Methods in Applied Econometrics: an Extended Analysis. Working paper, Michigan State University, Department of Economics.

# Supplementary material

## Appendix A: Literature review

We consider as dif-in-dif papers those where the authors explicitly describe their estimation strategy as dif-in-dif or where a policy intervention affects differently periods/groups and a dif-in-dif estimator is implicitly exploited. A paper is recorded as having a continuous outcome if at least one dependent variable is continuous (or discrete but with many mass points, such as hours worked). In cases where multiple outcomes were analysed, we report only the continuous ones or the one with several mass points. A paper is recorded as having a logged outcome if at least one dependent variable undergoes a log transformation.

Table A.1: Literature review: dif-in-dif papers with continuous outcome from the QJE 2001-2011

| Author and Year | Continuous Outcome Summary | Outcome Logged | Explicit comment on why log transformation |
|---|---|---|---|
| Autor (2001) | hourly wage | ✓ | |
| Donohue and Levitt (2001) | crime rate, arrest rate, arrest level | ✓ | ✓ |
| Donohue et al. (2002) | black/white ratio: teacher salaries, pupil/teacher ratio, term length | | |
| Bertrand et al. (2004) | weekly earnings | ✓ | |
| Finkelstein (2004) | new vaccine clinical trials | | |
| Morgan et al. (2004) | absolute deviation from conditional mean growth in: gross state product, employment, personal income, state GNP | | |
| Khwaja and Mian (2005) | default rate of firm, loan amount, export value, export value/total loans | ✓ | |
| Bailey (2006) | hours/weeks worked | | |
| Stevenson and Wolfers (2006) | homicide rate | | |
| Bandiera et al. (2007) | workers' productivity | ✓ | |
| Bleakley (2007) | schooling and earnings | ✓ | |
| DellaVigna and Kaplan (2007) | vote share and total vote cast | ✓ | ✓ |
| Field (2007) | weekly hours in labour force | ✓ | |
| Finkelstein (2007) | admissions, patient days, beds, payroll expenditures, and total expenditures | ✓ | ✓ |
| Donohue and Levitt (2008) | arrests | ✓ | |
| Foote and Goetz (2008) | arrests | ✓ | |
| Gruber and Hungerman (2008) | charitable giving | ✓ | |
| Verhoogen (2008) | difference over time of an estimated coefficient from OLS and IV regressions of log outcome on log domestic sales | ✓ | |
| Björkman and Svensson (2009) | equipment and waiting times | | |
| Jayachandran and Lleras-Muney (2009) | life expectancy, literacy, fertility (birth rate, # births, female pop. Aged 15-45) | ✓ | |
| Jensen and Oster (2009) | children enrolment rate | | |
| Hong and Kacperczyk (2010) | coverage and forecast bias | | |
| Hornbeck (2010) | acres of land, productivity, land value | ✓ | ✓ |
| Dittmar (2011) | city growth | ✓ | |
| Numn and Qian (2011) | total population | ✓ | ✓ |
| Number of dif-in-dif papers: | Total | 25 | |
| | Logging outcome | 16 | |
| | Comment on why logs | 5 | |
| | Addressing estimation issues | 1 | |

# Appendix B: Further details for the applied example

## B.1 The policy

EMA was rolled out nationally in the UK from September 2004. This followed on from evidence documenting the poor international performance of the UK in terms of post 16 education take-up (see for example Blanden, Gregg, and Machin, 2005). It provided resources to low income young people staying in post compulsory education, with the objective of increasing take-up rates of post 16 schooling. The programme was far reaching - by December of the 2010/11 academic year, 44% of 16-18 year olds in full-time education had received an EMA payment or 31% of all 16-18 year olds.[1]

EMA required attendance of a full time educational course at a school or college; or a course leading to an apprenticeship or foundation learning programme.[2] The policy was rolled out nationwide in September 2004. Therefore, students in low income households turning age 16 before that date and entering non-advanced education would be eligible for the allowance. The maximum value of the award was worth a considerable £30 per week, which was paid to the bank account of the eligible young person. Further bonus payments were available for meeting educational targets on course attendance, exam attendance and the completion of coursework assignments. These were paid twice annually and worth £100. It was usually claimed for two years of study, ending in the academic year a claimant turned 19. The final value of the award depended on household income and the meeting of educational targets, with the greatest amounts going to the poorest students.[3]

Prior to the national rollout, EMA was piloted in 15 Local Education Authorities

---

[1]Based on authors calculations from "Participation in Education, Training and Employment by 16-18 Year Olds in England , End 2011", Department for Education; and "Educational Maintenance Allowance Take-up", Young People's Learning Agency.

[2]A course is defined as one with at least 12 teaching hours per week, lasting for at least 10 weeks and be at an institution that is inspected by a public body, to assure quality.

[3]Income of the claimant was not included in the calculation of household income, however, claimants were restricted to engaging in part-time work of at most 24 hours per week. EMA did not interact with other UK benefits.

(LEAs), followed by a further 41 in the year 2000.[4] Therefore, for the national rollout, the pool of recipients consisted of those aged 16 at the start of the 2004/05 academic year (September), 16-17 year olds in 2005/06 academic year and 16-18 year olds in 2006/07 academic year; whilst in the pilot regions, EMA was rolled out from the 1999/2000 academic year. Dearden, Emmerson, and Meghir (2009) provide an excellent description of the policy environment and piloting of the programme. Their evidence suggests that the piloting of the scheme had substantial policy effects, with estimates implying a 4.5 percentage point increase in the first year of post 16 schooling and 6.7 percentage points in the second year. Further results indicate that the increased school participation largely comes from those who are not otherwise working and have low prior ability levels.

## B.2  Data details

The EFS is managed by the Office for National Statistics. The data is available online through the Economic and Social Data Service. The survey changed in 2008 to become the Living Costs and Food Survey. The primary purpose of the EFS is to provide expenditure weights for the consumer and retail price indexes. The survey records all expenditure items for a random sample of UK households. Expenditure items for all individuals aged over 7 in a household are recorded through a detailed expenditure diary over a two week period. Expenditures are then aggregated to the household level and into broad expenditure categories, and finally converted in weekly equivalents. The survey thus provides household level expenditure information for broad expenditure categories and disaggregated expenditures on specific consumption items.

---

[4]LEAs are local authorities responsible for education. There are 152 LEAs in England and Wales.

Table B.1: Summary statistics

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| **Household Characteristics** | | | | | |
| Number aged 16-18 | 2631 | 0.476245 | 0.622843 | 0 | 4 |
| HH Labour Income (Less Income from 14-18 year olds) | 2631 | 365.9028 | 390.1243 | 0 | 7028.37 |
| HH Size | 2631 | 4.026226 | 1.228804 | 1 | 12 |
| HH Owned | 2631 | 0.694793 | 0.460583 | 0 | 1 |
| Social Housing | 2631 | 0.22805 | 0.419655 | 0 | 1 |
| *Region:* | | | | | |
| North East | 2631 | 0.04333 | 0.203637 | 0 | 1 |
| North West | 2631 | 0.08932 | 0.285259 | 0 | 1 |
| Merseyside | 2631 | 0.027366 | 0.163179 | 0 | 1 |
| Yorkshire and Humber | 2631 | 0.082478 | 0.275144 | 0 | 1 |
| East Midlands | 2631 | 0.065755 | 0.2479 | 0 | 1 |
| West Midlands | 2631 | 0.083618 | 0.276867 | 0 | 1 |
| Eastern | 2631 | 0.08894 | 0.284711 | 0 | 1 |
| London | 2631 | 0.09274 | 0.290123 | 0 | 1 |
| South East | 2631 | 0.122767 | 0.328232 | 0 | 1 |
| South West | 2631 | 0.079818 | 0.271062 | 0 | 1 |
| Wales | 2631 | 0.049031 | 0.215973 | 0 | 1 |
| Scotland | 2631 | 0.076777 | 0.266288 | 0 | 1 |
| Northern Ireland | 2631 | 0.098062 | 0.297454 | 0 | 1 |
| **Expenditures** | | | | | |
| Food non-alcoholic drinks and Clothing | 2631 | 67.87192 | 34.07718 | 0 | 238.9919 |
| Alcoholic Beverages and Tobacco | 2631 | 16.26411 | 21.31974 | 0 | 196.0791 |
| Clothing and Footwear | 2631 | 42.43223 | 49.9206 | 0 | 526.4247 |
| Furnishings, HH Equipment, Carpets | 2631 | 40.58851 | 85.798 | 0 | 2017.65 |
| Transport | 2631 | 82.40274 | 110.0538 | 0 | 1310.762 |
| Communication | 2631 | 17.54936 | 15.77693 | 0 | 172.0529 |
| Recreation | 2631 | 85.82816 | 100.6262 | 0 | 1188.092 |
| **Observations** | | | | | |
| Treated | 1090 | - | - | - | - |
| Control | 1541 | - | - | - | - |
| NOTE: Summary statistics for full sample of treatment and control groups | | | | | |

Table B.2: Pre-reform summary statistics

|  | Treated | Control | Mean Diff |
|---|---|---|---|
| **Household Characteristics** | | | |
| Number aged 16-18 | 1.15 | 0.00 | 1.15*** |
| HH Labour Income (Less Income from 14-18 year olds) | 362.90 | 353.35 | 9.55 |
| HH Size | 4.09 | 4.00 | 0.10 |
| HH Owned | 0.69 | 0.69 | 0.00 |
| Social Housing | 0.22 | 0.25 | -0.03 |
| North East | 0.05 | 0.04 | 0.01 |
| North West | 0.09 | 0.09 | 0.00 |
| Merseyside | 0.03 | 0.03 | -0.00 |
| Yorkshire and the Humber | 0.08 | 0.08 | -0.00 |
| East Midlands | 0.06 | 0.07 | -0.00 |
| West Midlands | 0.08 | 0.10 | -0.02 |
| Eastern | 0.08 | 0.10 | -0.01 |
| London | 0.09 | 0.09 | 0.01 |
| South East | 0.11 | 0.12 | -0.01 |
| South West | 0.08 | 0.07 | 0.01 |
| Wales | 0.06 | 0.05 | 0.01 |
| Scotland | 0.08 | 0.08 | 0.00 |
| Northern Ireland | 0.11 | 0.10 | 0.01 |
| **Expenditures** | | | |
| Food non-alcoholic drinks and Clothing | 69.31 | 66.32 | 2.99 |
| Alcoholic Beverages and Tobacco | 18.47 | 15.83 | 2.65* |
| Clothing and Footwear | 50.55 | 38.54 | 12.01*** |
| Furnishings, HH Equipment, Carpets | 40.51 | 40.26 | 0.25 |
| Transport | 90.11 | 78.83 | 11.28* |
| Communication | 19.88 | 15.81 | 4.08*** |
| Recreation | 86.97 | 87.95 | -0.98 |
| Observations | 726 | 1073 | |

NOTE: * p<.10, ** p<.05, *** p<.01. Control group formed of households with at least one individual aged 14-15 (excluding households with 16-18 year olds). All expenditure categories are in weekly equivalent and are expressed in December 2005 terms.

## B.3 Possible violations of dif-in-dif assumptions

One of the possible problems in giving a causal interpretation to dif-in-dif estimates is that the treatment group assignment could reflect short-term idiosyncratic shocks, causing an Ashenfelter's dip (see Ashenfelter (1978) and Blundell and Dias (2009)). In such a scenario units assigned to the treatment group may recover more quickly in terms of the outcome of interest, than those in the control group. This is not a likely problem for the identification strategy outlined above, where treatment group assignment is allocated according to information on age.

Another possible source of bias is the presence of "anticipation effects", where households changed their spending behaviour prior to the reform. Pre-reform control households with younger children could potentially adjust their spending behaviour in anticipation of becoming EMA eligible in the post-reform period. This would lead to a downward bias in the estimated reform effects, assuming control households anticipating eligibility increased current spending. For the pre-reform treatment group (i.e. those with a 17 year old member), there is no problem of anticipation. Given that the roll-out of the policy applied only to new entrants in post 16 education, these individuals were ineligible.[5] A final important issue that could affect the causal interpretation of the dif-in-dif estimates is the presence of general equilibrium effects that influenced the spending behaviour of the control group. EMA is a large programme and any increase in post 16 participation rates implies increased competition for post 16 schooling places. This may further affect the spending behaviour of the control households in the event that it caused a change in their expected post 16 schooling plans or future expected wage rates, which in turn lead them to adjust their current spending behaviour.[6]

---

[5]However, if there are younger siblings in the household then anticipation effects are theoretically possible.

[6]The models estimated in this section do not include household level covariates. In a model with covariates, consistent estimation of the treatment effect further requires that the covariates are exogenous to the reform.

## B.4 A note on rescaling

As discussed in the methodology section, the initial piloting of the scheme means that in 41 of the 150 English LEAs, EMA was in operation before the start of the sample period. The above estimates therefore reflect a lower bound for the effect of EMA on the treated. Whilst the EFS data does not record information on LEA status, Government Office Region (GOR) information is available with each GOR being made up of multiple LEAs. Given information on EMA receipt by LEA, one may wonder whether it is possible to rescale the estimated EMA effects to reflect the fact that treatment on the treated group is less than 100%, but by a known number. Here, we point out that rescaling a multiplicative treatment effect may not always make sense.[7]

We know that when effects are heterogeneous, PPML returns the multiplicative effect on the average. For the case of EMA, we know that the effect of the treatment is zero for a known share $p$ of recipients. Following the main text, $exp(\delta)-1$ identifies

$$\frac{E\left[y_{1igt}|g=treated\right]-E\left[y_{0igt}|g=treated\right]}{E\left[y_{0igt}|g=treated\right]}. \tag{1}$$

However, both the numerator and denominator are going to be a weighted average of the two groups counterfactuals, so that $exp(\delta)-1$ is equal to

$$\frac{(E\left[y_{1it}|g=treated,LEA=notpilot\right]-E\left[y_{0igt}|g=treated,LEA=notpilot\right])\times(1-p)}{E\left[y_{0igt}|g=treated,LEA=pilot\right]\times p+E\left[y_{0igt}|g=treated,LEA=notpilot\right]\times(1-p)}. \tag{2}$$

Under the assumption that

$$E\left[y_{0igt}|g=treated,LEA=pilot\right]=E\left[y_{0igt}|g=treated,LEA=notpilot\right] \tag{3}$$

it is fairly trivial to rescale the estimated multiplicative effects of the results section.

---

[7]On the one hand, one might argue that as a policy maker it is enough to know the intent to treat.

The scale factor $(1 - p)$ could be calculated from publicly available data on EMA receipt by GOR and by then appropriately weighting for regional shares from the main estimation sample. However, eq. 3 is unlikely to hold in this example, where the pilot regions were on average much poorer than the national rollout areas. For this reason, we argue that rescaling may not make sense and caution against making such adjustments to the intent to treat estimates.

# Appendix C: Simulation of a reform with only mean effects, in the presence of a different variance across groups

We consider a separate scenario corresponding to the case where $\alpha = 0$, so that the multiplicative error term is statistically independent of the treatment, but where its properties depend on the group. Specifically, $y_{it}$ is heteroskedastic with respect to the group status but not to the treatment itself. To illustrate this scenario, consider the simulation procedure above but now $\sigma_{it}$ is generated according to:

$$\sigma_{it}^2 = exp(\gamma \times \mathbf{1}(treated_{it} = 1))$$

where $\gamma$ is now the parameter determining the degree of heteroskedasticity with respect to group status.

Simulations were made for five values of $\gamma = \{0, 0.1, 0.2, 0.3, 0.4\}$. Estimates of the multiplicative treatment effects along with the multiplicative treated group effects from OLS Log and PPML estimators are presented in table C.1, along with the OLS level models.

The table confirms that, under the particular form of heteroskedasticity considered, log linearisation works well when estimating the multiplicative treatment effect, but it performs poorly in terms of estimating the treated group effect. In

Table C.1: Simulation

| | (1) OLS Log | (2) PPML | (3) OLS Level |
|---|---|---|---|
| $\gamma = 0$ | | | |
| Treated x Post | .1976176 | .19624 | 4.618121 |
| | (0.0698541) | (0.084088) | (2.510334) |
| Treated | -.3995088 | -.3996061 | -10.8962 |
| | (0.0404981) | (0.047648) | (1.302451) |
| Park Test | 0.053 | 0.008 | - |
| Breusch-Pagan | 0.061 | - | - |
| $\gamma = 0.1$ | | | |
| Treated x Post | .1976049 | .1960462 | 4.614949 |
| | (0.0712667) | (0.0866465) | (2.570692) |
| Treated | -.4251419 | -.3996628 | -10.89587 |
| | (0.0413711) | (0.0490935) | (1.328722) |
| Park Test | 0.096 | 0.023 | - |
| Breusch-Pagan | 0.255 | - | - |
| $\gamma = 0.2$ | | | |
| Treated x Post | .197592 | .1958383 | 4.61164 |
| | (0.0727196) | (0.0893964) | (2.636176) |
| Treated | -.4520214 | -.3997227 | -10.89545 |
| | (0.0422676) | (0.0506476) | (1.357337) |
| Park Test | 0.242 | 0.063 | - |
| Breusch-Pagan | 0.705 | - | - |
| $\gamma = 0.4$ | | | |
| Treated x Post | .1975658 | .1953774 | 4.604632 |
| | (0.0757344) | (0.0955096) | (2.783938) |
| Treated | -.5094768 | -.3998539 | -10.89427 |
| | (0.0441245) | (0.0541067) | (1.422332) |
| Park Test | 0.715 | 0.304 | - |
| Breusch-Pagan | 0.996 | - | - |

NOTE: Results from 10,000 replications of the simulation procedure described in section 4. Mean of the estimated coefficients reported with standard deviations in parentheses.

contrast, PPML performs well for both the treated group and *treated × post* coefficients and for all values of $\gamma$. This illustrates the point made in section 2 that, in the dif-in-dif setting, heteroskedasticity is only a problem for consistently estimating the treatment effect when the treatment itself (and not the group) has an independent distributional effect; although estimates of the treated group effect can be misleading.

Finally, for the wrongly specified OLS level model in column 3, estimates of the multiplicative treatment effect again confound treatment and trend effects. Additionally, here we see that the form of heteroskedasticity means that the estimates of the treated group effect become more dispersed for higher values of gamma.

## Additional tables and figures
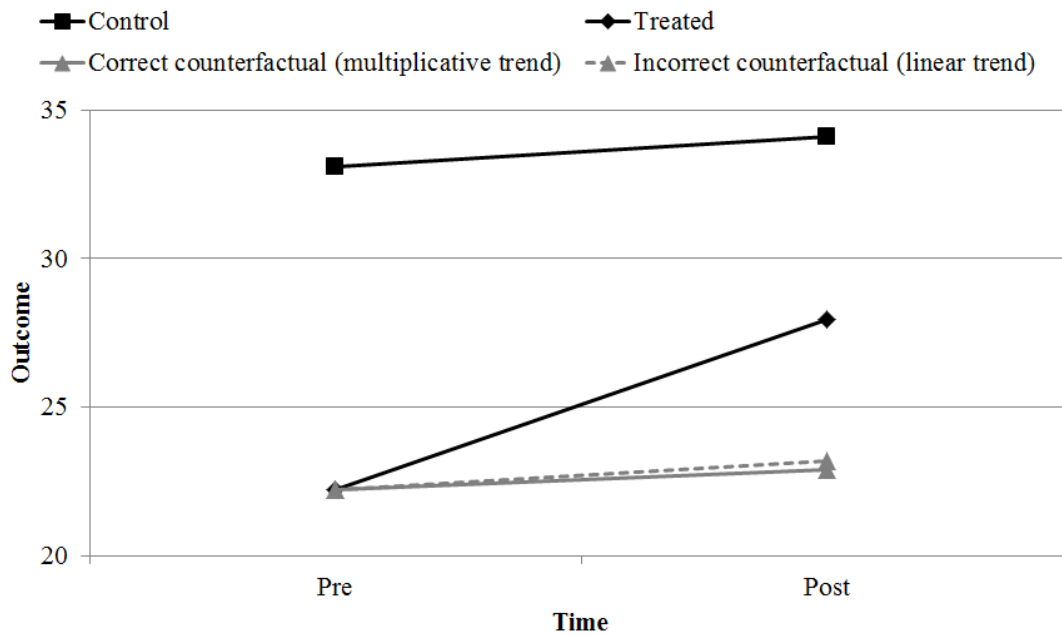


Figure C.1: Outcomes and counterfactuals in the simulation. The figure shows the expected value of the actual and counterfactual outcome for the different groups according to the exponential model $y = exp\left(\beta_0 + \beta_1 treated_{it} + \beta_2 post_{it} + \delta treated_{it} \times post_{it}\right)\eta_{it}$ with $\beta_0 = 3.5$, $\beta_1 = -0.4$, $\beta_2 = 0.03$, $\delta = 0.2$.

Table C.2: Summary of estimation methods and interpretation of the effect

| | Time trend | |
| | Additive | Multiplicative |
|---|---|---|
| No heteroskedasticity | OLS on model in levels; average additive effect on the treated | OLS on log-linearised model; multiplicative effect on the average for the treated |
| Heteroskedasticity by group and time | OLS model in levels, robust s.e.; average additive effect on the treated | PPML on exponential model, robust s.e.; multiplicative effect on the average for the treated |

# Additional references

Ashenfelter, O. (1978). Estimating the Effect of Training Programs on Earnings. *The Review of Economics and Statistics 60*, 47–57.

Autor, D. H. (2001). Why do temporary help firms provide free general skills training? *The Quarterly Journal of Economics 116*, 1409–1448.

Bailey, M. J. (2006). More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply. *The Quarterly Journal of Economics 121*, 289–320.

Bandiera, O., I. Barankay, and I. Rasul (2007). Incentives for managers and inequality among workers: Evidence from a firm-level experiment. *The Quarterly Journal of Economics 122*, 729–773.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*, 249–275.

Björkman, M. and J. Svensson (2009). Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda. *The Quarterly Journal of Economics 124*, 735–769.

Blanden, J., P. Gregg, and S. Machin (2005). Educational inequality and intergenerational mobility. In S. Machin and A. Vignoles (Eds.), *What's the good of education? The economics of education in the United Kingdom, Princeton University Press*. Princeton: Princeton University Press.

Bleakley, H. (2007). Disease and development: Evidence from hookworm eradication in the American South. *The Quarterly Journal of Economics 122*, 73–117.

Blundell, R. and M. C. Dias (2009). Alternative Approaches to Evaluation in Empirical Microeconomics. *Journal of Human Resources 44*, 565–640.

Dearden, L., C. Emmerson, and C. Meghir (2009). Conditional cash transfers and school dropout rates. *Journal of Human Resources 44*.

DellaVigna, S. and E. Kaplan (2007). The fox news effect: Media bias and voting. *The Quarterly Journal of Economics 122*, 1187–1234.

Dittmar, J. E. (2011). Information technology and economic change: The impact of the printing press. *The Quarterly Journal of Economics 126*, 1133–1172.

Donohue, J. J., J. J. Heckman, and P. E. Todd (2002). The Schooling of Southern Blacks: The Roles of Legal Activism and Private Philanthropy. *The Quarterly Journal of Economics 117*, 225–268.

Donohue, J. J. and S. D. Levitt (2001). The Impact of Legalised Abortion on Crime. *The Quarterly Journal of Economics 116*, 379–420.

Donohue, J. J. and S. D. Levitt (2008). Measurement error, legalized abortion, and the Decline in Crime: A Response to Foote and Goetz. *The Quarterly Journal of Economics 123*, 425–440.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in Peru. *The Quarterly Journal of Economics 122*, 1561–1602.

Finkelstein, A. (2004). Static and dynamic effects of health policy: Evidence from the vaccine industry. *The Quarterly Journal of Economics 119*, 527–564.

Finkelstein, A. (2007). The aggregate effects of health insurance: Evidence from the introduction of medicare. *The Quarterly Journal of Economics 122*, 1–37.

Foote, C. L. and C. F. Goetz (2008). The impact of legalized abortion on crime: Comment. *The Quarterly Journal of Economics 123*, 407–423.

Gruber, J. and D. M. Hungerman (2008). The church versus the mall: What happens when religion faces increased secular competition? *The Quarterly Journal of Economics 123*, 831–862.

Hong, H. and M. Kacperczyk (2010). Competition and bias. *The Quarterly Journal of Economics 125*, 1683–1725.

Hornbeck, R. (2010). Barbed wire: Property rights and agricultural development. *The Quarterly Journal of Economics 125*, 767–810.

Jayachandran, S. and A. Lleras-Muney (2009). Life expectancy and human capital investments: Evidence from maternal mortality declines. *The Quarterly Journal of Economics 124*, 349–397.

Jensen, R. and E. Oster (2009). The power of tv: Cable television and women's status in India. *The Quarterly Journal of Economics 124*, 1057–1094.

Khwaja, A. I. and A. Mian (2005). Do lenders favor politically connected firms? Rent provision in an emerging financial market. *The Quarterly Journal of Economics 120*, 1371–1411.

Morgan, D. P., B. Rime, and P. E. Strahan (2004). Bank integration and state business cycles. *The Quarterly Journal of Economics 119*, 1555–1584.

Nunn, N. and N. Qian (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *The Quarterly Journal of Economics 126*, 593–650.

Stevenson, B. and J. Wolfers (2006). Bargaining in the shadow of the law: Divorce laws and family distress. *The Quarterly Journal of Economics 121*, 267–288.

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector. *The Quarterly Journal of Economics 123*, 489–530.

RECENTLY PUBLISHED "TEMI" (*)

N. 954 – *Two EGARCH models and one fat tail*, by Michele Caivano and Andrew Harvey (March 2014).

N. 955 – *My parents taught me. Evidence on the family transmission of values*, by Giuseppe Albanese, Guido de Blasio and Paolo Sestito (March 2014).

N. 956 – *Political selection in the skilled city*, by Antonio Accetturo (March 2014).

N. 957 – *Calibrating the Italian smile with time-varying volatility and heavy-tailed models*, by Michele Leonardo Bianchi (April 2014).

N. 958 – *The intergenerational transmission of reading: is a good example the best sermon?*, by Anna Laura Mancini, Chiara Monfardini and Silvia Pasqua (April 2014).

N. 959 – *A tale of an unwanted outcome: transfers and local endowments of trust and cooperation*, by Antonio Accetturo, Guido de Blasio and Lorenzo Ricci (April 2014).

N. 960 – *The impact of R&D subsidies on firm innovation*, by Raffaello Bronzini and Paolo Piselli (April 2014).

N. 961 – *Public expenditure distribution, voting, and growth*, by Lorenzo Burlon (April 2014).

N. 962 – *Cooperative R&D networks among firms and public research institutions*, by Marco Marinucci (June 2014).

N. 963 – *Technical progress, retraining cost and early retirement*, by Lorenzo Burlon and Montserrat Vilalta-Bufí (June 2014).

N. 964 – Foreign exchange reserve diversification and the "exorbitant privilege", by Pietro Cova, Patrizio Pagano and Massimiliano Pisani (July 2014).

N. 965 – *Behind and beyond the (headcount) employment rate*, by Andrea Brandolini and Eliana Viviano (July 2014).

N. 966 – *Bank bonds: size, systemic relevance and the sovereign*, by Andrea Zaghini (July 2014).

N. 967 – *Measuring spatial effects in presence of institutional constraints: the case of Italian Local Health Authority expenditure*, by Vincenzo Atella, Federico Belotti, Domenico Depalo and Andrea Piano Mortari (July 2014).

N. 968 – *Price pressures in the UK index-linked market: an empirical investigation*, by Gabriele Zinna (July 2014).

N. 969 – *Stock market efficiency in China: evidence from the split-share reform*, by Andrea Beltratti, Bernardo Bortolotti and Marianna Caccavaio (September 2014).

N. 970 – *Academic performance and the Great Recession*, by Effrosyni Adamopoulou and Giulia Martina Tanzi (September 2014).

N. 971 – *Random switching exponential smoothing and inventory forecasting*, by Giacomo Sbrana and Andrea Silvestrini (September 2014).

N. 972 – *Are Sovereign Wealth Funds contrarian investors?*, by Alessio Ciarlone and Valeria Miceli (September 2014).

N. 973 – *Inequality and trust: new evidence from panel data*, by Guglielmo Barone and Sauro Mocetti (September 2014).

N. 974 – *Identification and estimation of outcome response with heterogeneous treatment externalities*, by Tiziano Arduini, Eleonora Patacchini and Edoardo Rainone (September 2014).

N. 975 – *Hedonic value of Italian tourism supply: comparing environmental and cultural attractiveness*, by Valter Di Giacinto and Giacinto Micucci (September 2014).

N. 976 – *Multidimensional poverty and inequality*, by Rolf Aaberge and Andrea Brandolini (September 2014).

*2011*

S. DI ADDARIO, *Job search in thick markets*, Journal of Urban Economics, v. 69, 3, pp. 303-318, **TD No. 605 (December 2006).**

F. SCHIVARDI and E. VIVIANO, *Entry barriers in retail trade*, Economic Journal, v. 121, 551, pp. 145-170, **TD No. 616 (February 2007).**

G. FERRERO, A. NOBILI and P. PASSIGLIA, *Assessing excess liquidity in the Euro Area: the role of sectoral distribution of money,* Applied Economics, v. 43, 23, pp. 3213-3230, **TD No. 627 (April 2007).**

P. E. MISTRULLI, *Assessing financial contagion in the interbank market: maximun entropy versus observed interbank lending patterns,* Journal of Banking & Finance, v. 35, 5, pp. 1114-1127, **TD No. 641 (September 2007).**

E. CIAPANNA, *Directed matching with endogenous markov probability: clients or competitors?*, The RAND Journal of Economics, v. 42, 1, pp. 92-120, **TD No. 665 (April 2008).**

M. BUGAMELLI and F. PATERNÒ, *Output growth volatility and remittances*, Economica, v. 78, 311, pp. 480-500, **TD No. 673 (June 2008).**

V. DI GIACINTO e M. PAGNINI, *Local and global agglomeration patterns: two econometrics-based indicators,* Regional Science and Urban Economics, v. 41, 3, pp. 266-280, **TD No. 674 (June 2008).**

G. BARONE and F. CINGANO, *Service regulation and growth: evidence from OECD countries,* Economic Journal, v. 121, 555, pp. 931-957, **TD No. 675 (June 2008).**

P. SESTITO and E. VIVIANO, *Reservation wages: explaining some puzzling regional patterns*, Labour, v. 25, 1, pp. 63-88, **TD No. 696 (December 2008).**

R. GIORDANO and P. TOMMASINO, *What determines debt intolerance? The role of political and monetary institutions*, European Journal of Political Economy, v. 27, 3, pp. 471-484, **TD No. 700 (January 2009).**

P. ANGELINI, A. NOBILI and C. PICILLO, *The interbank market after August 2007: What has changed, and why?*, Journal of Money, Credit and Banking, v. 43, 5, pp. 923-958, **TD No. 731 (October 2009).**

G. BARONE and S. MOCETTI, *Tax morale and public spending inefficiency,* International Tax and Public Finance, v. 18, 6, pp. 724-49, **TD No. 732 (November 2009).**

L. FORNI, A. GERALI and M. PISANI, *The Macroeconomics of Fiscal Consolidation in a Monetary Union: the Case of Italy,* in Luigi Paganetto (ed.), Recovery after the crisis. Perspectives and policies, VDM Verlag Dr. Muller, **TD No. 747 (March 2010).**

A. DI CESARE and G. GUAZZAROTTI, *An analysis of the determinants of credit default swap changes before and during the subprime financial turmoil*, in Barbara L. Campos and Janet P. Wilkins (eds.), The Financial Crisis: Issues in Business, Finance and Global Economics, New York, Nova Science Publishers, Inc., **TD No. 749 (March 2010).**

A. LEVY and A. ZAGHINI, *The pricing of government guaranteed bank bonds,* Banks and Bank Systems, v. 6, 3, pp. 16-24, **TD No. 753 (March 2010).**

G. BARONE, R. FELICI and M. PAGNINI, *Switching costs in local credit markets,* International Journal of Industrial Organization, v. 29, 6, pp. 694-704, **TD No. 760 (June 2010).**

G. BARBIERI, C. ROSSETTI e P. SESTITO, *The determinants of teacher mobility: evidence using Italian teachers' transfer applications*, Economics of Education Review, v. 30, 6, pp. 1430-1444, **TD No. 761 (marzo 2010).**

G. GRANDE and I. VISCO, *A public guarantee of a minimum return to defined contribution pension scheme members,* The Journal of Risk, v. 13, 3, pp. 3-43, **TD No. 762 (June 2010).**

P. DEL GIOVANE, G. ERAMO and A. NOBILI, *Disentangling demand and supply in credit developments: a survey-based analysis for Italy,* Journal of Banking and Finance, v. 35, 10, pp. 2719-2732, **TD No. 764 (June 2010).**

G. BARONE and S. MOCETTI, *With a little help from abroad: the effect of low-skilled immigration on the female labour supply,* Labour Economics, v. 18, 5, pp. 664-675, **TD No. 766 (July 2010).**

S. FEDERICO and A. FELETTIGH, *Measuring the price elasticity of import demand in the destination markets of italian exports,* Economia e Politica Industriale, v. 38, 1, pp. 127-162, **TD No. 776 (October 2010).**

S. MAGRI and R. PICO, *The rise of risk-based pricing of mortgage interest rates in Italy*, Journal of Banking and Finance, v. 35, 5, pp. 1277-1290, **TD No. 778 (October 2010).**

M. TABOGA, *Under/over-valuation of the stock market and cyclically adjusted earnings*, International Finance, v. 14, 1, pp. 135-164, **TD No. 780 (December 2010).**

S. NERI, *Housing, consumption and monetary policy: how different are the U.S. and the Euro area?*, Journal of Banking and Finance, v.35, 11, pp. 3019-3041, **TD No. 807 (April 2011).**

V. CUCINIELLO, *The welfare effect of foreign monetary conservatism with non-atomistic wage setters*, Journal of Money, Credit and Banking, v. 43, 8, pp. 1719-1734, **TD No. 810 (June 2011).**

A. CALZA and A. ZAGHINI, *welfare costs of inflation and the circulation of US currency abroad,* The B.E. Journal of Macroeconomics, v. 11, 1, Art. 12, **TD No. 812 (June 2011).**

I. FAIELLA, *La spesa energetica delle famiglie italiane*, Energia, v. 32, 4, pp. 40-46, **TD No. 822 (September 2011).**

D. DEPALO and R. GIORDANO, *The public-private pay gap: a robust quantile approach*, Giornale degli Economisti e Annali di Economia, v. 70, 1, pp. 25-64, **TD No. 824 (September 2011).**

R. DE BONIS and A. SILVESTRINI, *The effects of financial and real wealth on consumption: new evidence from OECD countries,* Applied Financial Economics, v. 21, 5, pp. 409–425, **TD No. 837 (November 2011).**

F. CAPRIOLI, P. RIZZA and P. TOMMASINO, *Optimal fiscal policy when agents fear government default*, Revue Economique, v. 62, 6, pp. 1031-1043, **TD No. 859 (March 2012).**


*2012*


F. CINGANO and A. ROSOLIA, *People I know: job search and social networks*, Journal of Labor Economics, v. 30, 2, pp. 291-332,  **TD No. 600 (September 2006).**

G. GOBBI and R. ZIZZA, *Does the underground economy hold back financial deepening? Evidence from the italian credit market*, Economia Marche, Review of Regional Studies, v. 31, 1, pp. 1-29, **TD No. 646 (November 2006).**

S. MOCETTI, *Educational choices and the selection process before and after compulsory school*, Education Economics, v. 20, 2, pp. 189-209, **TD No. 691 (September 2008).**

P. PINOTTI, M. BIANCHI and P. BUONANNO, *Do immigrants cause crime?*, Journal of the European Economic Association , v. 10, 6, pp. 1318–1347, **TD No. 698 (December 2008).**

M. PERICOLI and M. TABOGA, *Bond risk premia, macroeconomic fundamentals and the exchange rate*, International Review of Economics and Finance, v. 22, 1, pp. 42-65, **TD No. 699 (January 2009).**

F. LIPPI and A. NOBILI, *Oil and the macroeconomy: a quantitative structural analysis,* Journal of European Economic Association, v. 10, 5, pp. 1059-1083, **TD No. 704 (March 2009).**

G. ASCARI and T. ROPELE, *Disinflation in a DSGE perspective: sacrifice ratio or welfare gain ratio?,* Journal of Economic Dynamics and Control, v. 36, 2, pp. 169-182, **TD No. 736 (January 2010).**

S. FEDERICO, *Headquarter intensity and the choice between outsourcing versus integration at home or abroad*, Industrial and Corporate Chang, v. 21, 6, pp. 1337-1358, **TD No. 742 (February 2010).**

I. BUONO and G. LALANNE, *The effect of the Uruguay Round on the intensive and extensive margins of trade,* Journal of International Economics, v. 86, 2, pp. 269-283,  **TD No. 743 (February 2010).**

A. BRANDOLINI, S. MAGRI and T. M SMEEDING, *Asset-based measurement of poverty*, In D. J. Besharov and K. A. Couch (eds), Counting the Poor: New Thinking About European Poverty Measures and Lessons for the United States, Oxford and New York: Oxford University Press, **TD No. 755 (March 2010).**

S. GOMES, P. JACQUINOT and M. PISANI, *The EAGLE. A model for policy analysis of macroeconomic interdependence in the euro area,* Economic Modelling, v. 29, 5, pp. 1686-1714, **TD No. 770 (July 2010).**

A. ACCETTURO and G. DE BLASIO, *Policies for local development: an evaluation of Italy's "Patti Territoriali",* Regional Science and Urban Economics, v. 42, 1-2, pp. 15-26, **TD No. 789 (January 2006).**

F. BUSETTI and S. DI SANZO, *Bootstrap LR tests of stationarity, common trends and cointegration,* Journal of Statistical Computation and Simulation, v. 82, 9, pp. 1343-1355, **TD No. 799 (March 2006).**

S. NERI and T. ROPELE, *Imperfect information, real-time data and monetary policy in the Euro area,* The Economic Journal, v. 122, 561, pp. 651-674,  **TD No. 802 (March 2011).**

A. ANZUINI and F. FORNARI, *Macroeconomic determinants of carry trade activity,* Review of International Economics, v. 20, 3, pp. 468-488,  **TD No. 817 (September 2011).**

M. AFFINITO, *Do interbank customer relationships exist? And how did they function in the crisis? Learning from Italy*, Journal of Banking and Finance, v. 36, 12, pp. 3163-3184, **TD No. 826 (October 2011).**

P. GUERRIERI and F. VERGARA CAFFARELLI, *Trade Openness and International Fragmentation of Production in the European Union: The New Divide?,* Review of International Economics, v. 20, 3, pp. 535-551,  **TD No. 855 (February 2012).**

V. DI GIACINTO, G. MICUCCI and P. MONTANARO, *Network effects of public transposrt infrastructure: evidence on Italian regions,* Papers in Regional Science, v. 91, 3, pp. 515-541, **TD No. 869 (July 2012).**

A. FILIPPIN and M. PACCAGNELLA, *Family background, self-confidence and economic outcomes,* Economics of Education Review, v. 31, 5, pp. 824-834,  **TD No. 875 (July 2012).**


*2013*


A. MERCATANTI, *A likelihood-based analysis for relaxing the exclusion restriction in randomized experiments with imperfect compliance,* Australian and New Zealand Journal of Statistics, v. 55, 2, pp. 129-153, **TD No. 683 (August 2008).**

F. CINGANO and P. PINOTTI, *Politicians at work. The private returns and social costs of political connections,* Journal of the European Economic Association, v. 11, 2, pp. 433-465, **TD No. 709 (May 2009).**

F. BUSETTI and J. MARCUCCI, *Comparing forecast accuracy: a Monte Carlo investigation,* International Journal of Forecasting, v. 29, 1, pp. 13-27, **TD No. 723 (September 2009).**

D. DOTTORI, S. I-LING and F. ESTEVAN, *Reshaping the schooling system: The role of immigration,* Journal of Economic Theory, v. 148, 5, pp. 2124-2149, **TD No. 726 (October 2009).**

A. FINICELLI, P. PAGANO and M. SBRACIA, *Ricardian Selection,* Journal of International Economics, v. 89, 1, pp. 96-109, **TD No. 728 (October 2009).**

L. MONTEFORTE and G. MORETTI, *Real-time forecasts of inflation: the role of financial variables,* Journal of Forecasting,  v. 32,  1, pp. 51-61, **TD No. 767 (July 2010).**

R. GIORDANO and P. TOMMASINO, *Public-sector efficiency and political culture,* FinanzArchiv, v. 69, 3, pp. 289-316, **TD No. 786 (January 2011).**

E. GAIOTTI, *Credit availablility and investment: lessons from the "Great Recession",* European Economic Review, v. 59, pp. 212-227, **TD No. 793 (February 2011).**

F. NUCCI and M. RIGGI, *Performance pay and changes in U.S. labor market dynamics,* Journal of Economic Dynamics and Control, v. 37, 12, pp. 2796-2813,  **TD No. 800 (March 2011).**

G. CAPPELLETTI, G. GUAZZAROTTI and P. TOMMASINO, *What determines annuity demand at retirement?,* The Geneva Papers on Risk and Insurance – Issues and Practice, pp. 1-26, **TD No. 805 (April 2011).**

A. ACCETTURO e L. INFANTE, *Skills or Culture? An analysis of the decision to work by immigrant women in Italy,* IZA Journal of Migration, v. 2, 2, pp. 1-21, **TD No. 815 (July 2011).**

A. DE SOCIO, *Squeezing liquidity in a "lemons market" or asking liquidity "on tap",* Journal of Banking and Finance, v. 27, 5, pp. 1340-1358, **TD No. 819 (September 2011).**

S. GOMES, P. JACQUINOT, M. MOHR and M. PISANI, *Structural reforms and macroeconomic performance in the euro area countries: a model-based assessment,* International Finance, v. 16, 1, pp. 23-44, **TD No. 830 (October 2011).**

G. BARONE and G. DE BLASIO, *Electoral rules and voter turnout,* International Review of Law and Economics, v. 36, 1, pp. 25-35, **TD No. 833 (November 2011).**

O. BLANCHARD and M. RIGGI, *Why are the 2000s so different from the 1970s? A structural interpretation of changes in the macroeconomic effects of oil prices,* Journal of the European Economic Association, v. 11, 5, pp. 1032-1052,  **TD No. 835 (November 2011).**

R. CRISTADORO and D. MARCONI, *Household savings in China,* in G. Gomel, D. Marconi, I. Musu, B. Quintieri (eds), The Chinese Economy: Recent Trends and Policy Issues, Springer-Verlag, Berlin, **TD No. 838 (November 2011).**

E. GENNARI and G. MESSINA, *How sticky are local expenditures in Italy? Assessing the relevance of the flypaper effect through municipal data,* International Tax and Public Finance (DOI: 10.1007/s10797-013-9269-9), **TD No. 844 (January 2012).**

A. ANZUINI, M. J. LOMBARDI and P. PAGANO, *The impact of monetary policy shocks on commodity prices,* International Journal of Central Banking, v. 9, 3, pp. 119-144, **TD No. 851 (February 2012).**

R. GAMBACORTA and M. IANNARIO, *Measuring job satisfaction with CUB models,* Labour, v. 27, 2, pp. 198-224,  **TD No. 852 (February 2012).**

G. ASCARI and T. ROPELE, *Disinflation effects in a medium-scale new keynesian model: money supply rule versus interest rate rule,* European Economic Review, v. 61, pp. 77-100, **TD No. 867 (April 2012).**

E. BERETTA and S. DEL PRETE, *Banking consolidation and bank-firm credit relationships: the role of geographical features and relationship characteristics,* Review of Economics and Institutions, v. 4, 3, pp. 1-46, **TD No. 901 (February 2013).**

M. ANDINI, G. DE BLASIO, G. DURANTON and W. STRANGE, *Marshallian labor market pooling: evidence from Italy,* Regional Science and Urban Economics, v. 43, 6, pp.1008-1022, **TD No. 922 (July 2013).**

G. SBRANA and A. SILVESTRINI, *Forecasting aggregate demand: analytical comparison of top-down and bottom-up approaches in a multivariate exponential smoothing framework,* International Journal of Production Economics, v. 146, 1, pp. 185-98, **TD No. 929 (September 2013).**

A. FILIPPIN, C. V, FIORIO and E. VIVIANO, *The effect of tax enforcement on tax morale,* European Journal of Political Economy, v. 32, pp. 320-331, **TD No. 937 (October 2013).**


*2014*


M. TABOGA, *The riskiness of corporate bonds*, Journal of Money, Credit and Banking, v.46, 4, pp. 693-713, **TD No. 730 (October 2009).**

G. MICUCCI and P. ROSSI, *Il ruolo delle tecnologie di prestito nella ristrutturazione dei debiti delle imprese in crisi*, in A. Zazzaro (a cura di), Le banche e il credito alle imprese durante la crisi, Bologna, Il Mulino, **TD No. 763 (June 2010).**

P. ANGELINI, S. NERI and F. PANETTA, *The interaction between capital requirements and monetary policy*, Journal of Money, Credit and Banking, v. 46, 6, pp. 1073-1112, **TD No. 801 (March 2011).**

M. FRANCESE and R. MARZIA, *Is there Room for containing healthcare costs? An analysis of regional spending differentials in Italy,* The European Journal of Health Economics, v. 15, 2, pp. 117-132, **TD No. 828 (October 2011).**

L. GAMBACORTA and P. E. MISTRULLI, *Bank heterogeneity and interest rate setting: what lessons have we learned since Lehman Brothers?,* Journal of Money, Credit and Banking, v. 46, 4, pp. 753-778, **TD No. 829 (October 2011).**

M. PERICOLI, *Real term structure and inflation compensation in the euro area*, International Journal of Central Banking, v. 10, 1, pp. 1-42, **TD No. 841 (January 2012).**

V. DI GACINTO, M. GOMELLINI, G. MICUCCI and M. PAGNINI, *Mapping local productivity advantages in Italy: industrial districts, cities or both?*, Journal of Economic Geography, v. 14, pp. 365–394, **TD No. 850 (January 2012).**

A. ACCETTURO, F. MANARESI, S. MOCETTI and E. OLIVIERI, *Don't Stand so close to me: the urban impact of immigration,* Regional Science and Urban Economics, v. 45, pp. 45-56, **TD No. 866 (April 2012).**

S. FEDERICO, *Industry dynamics and competition from low-wage countries: evidence on Italy*, Oxford Bulletin of Economics and Statistics, v. 76, 3, pp. 389-410, **TD No. 879 (September 2012).**

F. D'AMURI and G. PERI, *Immigration, jobs and employment protection: evidence from Europe before and during the Great Recession,* Journal of the European Economic Association, v. 12, 2, pp. 432-464, **TD No. 886 (October 2012).**

M. TABOGA, *What is a prime bank? A euribor-OIS spread perspective,* International Finance, v. 17, 1, pp. 51-75, **TD No. 895 (January 2013).**

L. GAMBACORTA and F. M. SIGNORETTI, *Should monetary policy lean against the wind? An analysis based on a DSGE model with banking,* Journal of Economic Dynamics and Control, v. 43, pp. 146-74, **TD No. 921 (July 2013).**

U. ALBERTAZZI and M. BOTTERO, *Foreign bank lending: evidence from the global financial crisis,* Journal of International Economics, v. 92, 1, pp. 22-35, **TD No. 926 (July 2013).**

R. DE BONIS and A. SILVESTRINI, *The Italian financial cycle: 1861-2011,* Cliometrica, v.8, 3, pp. 301-334, **TD No. 936 (October 2013).**

D. PIANESELLI and A. ZAGHINI, *The cost of firms' debt financing and the global financial crisis,* Finance Research Letters, v. 11, 2, pp. 74-83, **TD No. 950 (February 2014).**

A. ZAGHINI, *Bank bonds: size, systemic relevance and the sovereign,* International Finance, v. 17, 2, pp. 161-183, **TD No. 966 (July 2014).**

*FORTHCOMING*

M. BUGAMELLI, S. FABIANI and E. SETTE, *The age of the dragon: the effect of imports from China on firm-level prices*, Journal of Money, Credit and Banking, **TD No. 737 (January 2010).**

F. D'AMURI, *Gli effetti della legge 133/2008 sulle assenze per malattia nel settore pubblico,* Rivista di Politica Economica, **TD No. 787 (January 2011).**

E. COCOZZA and P. PISELLI, *Testing for east-west contagion in the European banking sector during the financial crisis,* in R. Matoušek; D. Stavárek (eds.), Financial Integration in the European Union, Taylor & Francis, **TD No. 790 (February 2011).**

R. BRONZINI and E. IACHINI, *Are incentives for R&D effective? Evidence from a regression discontinuity approach,* American Economic Journal : Economic Policy, **TD No. 791 (February 2011).**

G. DE BLASIO, D. FANTINO and G. PELLEGRINI, *Evaluating the impact of innovation incentives: evidence from an unexpected shortage of funds,* Industrial and Corporate Change, **TD No. 792 (February 2011).**

A. DI CESARE, A. P. STORK and C. DE VRIES, *Risk measures for autocorrelated hedge fund returns,* Journal of Financial Econometrics, **TD No. 831 (October 2011).**

D. FANTINO, A. MORI and D. SCALISE, *Collaboration between firms and universities in Italy: the role of a firm's proximity to top-rated departments,* Rivista Italiana degli economisti, **TD No. 884 (October 2012).**

G. BARONE and S. MOCETTI, *Natural disasters, growth and institutions: a tale of two earthquakes,* Journal of Urban Economics, **TD No. 949 (January 2014).**