Differencing as a Consistency Test for the Within Estimator

Spierdijk, Laura; Wansbeek, Thomas

university of groningen
faculty of economics and business

2021006-EEF

Differencing as a Consistency Test for the Within Estimator

April 2021

Laura Spierdijk
Tom Wansbeek

SOM is the research institute of the Faculty of Economics & Business at the University of Groningen. SOM has six programmes:
- Economics, Econometrics and Finance
- Global Economics & Management
- Innovation & Organization
- Marketing
- Operations Management & Operations Research
- Organizational Behaviour

Research Institute SOM
Faculty of Economics & Business
University of Groningen

Visiting address:
Nettelbosje 2
9747 AE Groningen
The Netherlands

Postal address:
P.O. Box 800
9700 AV Groningen
The Netherlands

T +31 50 363 9090/7068/3815

www.rug.nl/feb/research

Differencing as a Consistency Test for the Within Estimator

Laura Spierdijk
University of Groningen, Faculty of Economics and Business, Department of Economics, Econometrics and Finance

Tom Wansbeek
University of Groningen, Faculty of Economics and Business, Department of Economics, Econometrics and Finance
t.j.wansbeek@rug.nl

Differencing as a Consistency Test for the Within Estimator

Laura Spierdijk[a,*], Tom Wansbeek[a]

[a] *University of Groningen, Faculty of Economics and Business, Department of Economics, Econometrics and Finance, P.O. Box 800, 9700 AV Groningen, The Netherlands.*

Abstract

The within estimator is commonly used to estimate the linear panel regression model. We exploit the differences between short- and long-differences estimators to construct a GMM-test for the exogeneity assumption underlying the within estimator. We find that this test is locally more powerful than a more generic GMM-test for exogeneity of the regressors. We use our GMM-test in the representation of a Wald test, which facilitates the economic interpretation and visualization of the test outcomes. We illustrate our approach in an application to U.S. banks' economies of scale.

JEL codes: C23, C52

Keywords: linear panel regression, within estimator, differences estimators, measurement error

Acknowledgements

The authors are grateful to Vasilis Sarafidis and other participants of the 25th International Panel Data Conference in Vilnius on July 4–5, 2019. Laura Spierdijk gratefully acknowledges financial support by a Vidi grant (452.11.007) in the 'Vernieuwingsimpuls' program of the Netherlands Organization for Scientific Research (NWO). Her work was also supported by the Netherlands Institute for Advanced Study in the Humanities and Social Sciences (NIAS-KNAW). The usual disclaimer applies.

*Corresponding author

Email addresses: `L.Spierdijk@rug.nl` (Laura Spierdijk), `T.J.Wansbeek@rug.nl` (Tom Wansbeek)

1. Introduction

Since the early days of econometrics, the within estimator has been widely used to estimate the linear panel regression model in the presence of individual effects correlated with the regressors (Mundlak, 1961; Mundlak and Hoch, 1965). Its consistency requires exogeneity of the regressors after removal of the individual effect. If there are any doubts about a particular covariate's exogeneity and one or more instrumental variables are available, it is common to run a Hausman test for exogeneity or a J-test for instrument validity (also known as Hansen-Sargan, GMM or overidentifying test). Both tests are based on prior suspicions about certain covariates and rely on the availability of instrumental variables. In the absence of prior information or instruments, it is still important to test the validity of the within estimator's exogeneity assumption. To our best knowledge, however, a more general specification test does not exist for this widely occurring case. The present study seeks to remedy this situation.

To that end, we combine the ideas from the time-series literature about specification tests with the insights of Griliches and Hausman (1986) to construct a test for the within estimator's exogeneity assumption. From the time-series literature, we take the idea to develop a specification test that exploits model transformation; see e.g. Plosser et al. (1982), Davidson et al. (1985), Breusch and Godfrey (1986), and Thursby (1989). These studies use a Hausman test to compare the OLS estimators obtained from differenced and undifferenced regression models. Under the null hypothesis of no misspecification, OLS yields a consistent and efficient estimator for the undifferenced model, while it produces a consistent but inefficient estimator for the differenced model. The power of the Hausman test arises from the difference in the estimators' probability limits under misspecification. We combine this idea with the insight of Griliches and Hausman (1986, p. 114) that misspecification may be present in the linear panel regression model if short- and long-differences estimators differ significantly.

Although Griliches and Hausman (1986) has eventually become the most frequently cited study about measurement error in econometrics, exploiting the patterns in short- and long-differences estimators has hardly ever been done in more than thirty years. In fact, we found only four panel data studies that do this (Levitt, 1998; Goolsbee, 2000; McKinnish, 2008; Bun et al., 2019). These studies link the patterns in short- and long-differences estimators to measurement error, but do not consider their relevance for detecting general misspecification. By contrast, we show that these patterns can detect violation of the within estimator's exogeneity assumption due to misspecification in general.

Our specification test is developed in a GMM framework and has the familiar form of a J-test. We show that our test is locally more powerful than a J-test for all moment conditions that follow from the exogeneity of the regressors after removing the individual effects. We will refer to this alternative test as the 'generic' J-test, as opposed to our more specific 'differences' J-test. The explanation for the better performance of our J-test is that it directs the power from the full set of moment conditions for exogeneity to the moment conditions that are directly relevant for the within estimator. Throughout, we use our J-test in the representation of a Wald test, which facilitates the economic interpretation and visualization of the test outcomes. We also provide an empirical strategy for consistent model estimation based on our test.

In the empirical part of our analysis, we estimate U.S. banks' cost functions using the within estimator and calculate the implied scale elasticity during the 2011–2017 period. In the literature, estimates of banks scale elasticities have been used to plead against a size limit on banks (Hughes and Mester, 2013) and to determine the implied net costs of increasing bank size for too-big-to-fail banks (Boyd and Heitz, 2016). Our test finds strong evidence against the within estimator's consistency. We discuss the possible sources of endogeneity by relating our application to the long-standing problem in econometrics of how to consistently estimate cost and production functions. We also provide suggestions for further modeling.

The test developed in this study contributes to the panel-data literature about fixed-T and large-n specification testing, which includes but is not limited to tests for overidentifying restrictions (Hayakawa, 2019), random vs. fixed effects and for FE vs. FE-2SLS (Hausman, 1978; Baltagi et al., 2003; Amini et al., 2012; Joshi and Wooldridge, 2019), unit roots (Harris and Tzavalis, 1999), selectivity bias (Verbeek and Nijman, 1992; Wooldridge, 1995), cross-sectional dependence (Sarafidis and Wansbeek, 2012) and GMM-based test for autocorrelation in error terms (Arellano and Bond, 1991).

The setup of the remainder of this study is as follows. Section 2 describes the test statistic and discusses its statistical properties. Section 3 provides the empirical application to U.S. banks' scale elasticities. Lastly, Section 4 concludes. Proofs and additional results can be found in the appendix with supplementary material.

2. Test statistic

We consider the linear panel regression model with T observations over time, given by

$$\mathbf{y}_i = \gamma \boldsymbol{\iota}_T + \mathbf{X}_i \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i \qquad [i = 1, \ldots, n], \tag{1}$$

where $\mathbf{y}_i$ $(T \times 1)$ is the dependent variable, γ the intercept, $\boldsymbol{\iota}_T$ $(T \times 1)$ a vector of ones, $\mathbf{X}_i$ $(T \times k)$ the matrix of observed covariates, $\boldsymbol{\beta}$ $(k \times 1)$ the coefficient vector, and $\boldsymbol{\varepsilon}_i$ $(T \times 1)$ the error term containing an individual effect possibly correlated with $\mathbf{X}_i$.

The within estimator is widely used to estimate $\boldsymbol{\beta}$ in (1). However, its consistency requires exogeneity of $\mathbf{X}_i$, after removing the individual effects. This is a strong assumption, which can be violated due to e.g. measurement error, omitted variables or simultaneity. For this reason, we propose a test for the consistency of the within estimator.

2.1. *J-test*

The test we propose is developed in a GMM framework and of the familiar form of the J-test for some population moment condition $H_0 : \mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})] = 0$, exploiting overidentification.

The statistical properties of the J-test under correct specification and misspecification are well-known, but have been mostly studied in a time-series context (Newey, 1985; Hall, 2005). These properties rely on the asymptotic normality of the GMM estimator. We refer to Hayakawa (2019) for the formulation of similar conditions in a fixed-T and large-n panel data setting. Throughout, we assume that these conditions hold and that the usual properties of the J-test apply. In particular, we use that the asymptotic distribution of the J-test is central chi-square under H_0 and non-central chi-square under local alternatives of the Pitman form $H_1 : \mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})] = \mathbf{d}/\sqrt{n}$, for some finite constant $\mathbf{d}$.

We start with some notation. Let $\mathbf{e}_\ell$ $(k \times 1)$ be the ℓ-th unit vector and write

$$\mathbf{X}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{ik}) = \sum_\ell \mathbf{x}_{i\ell}\mathbf{e}'_\ell \qquad \text{such that} \qquad \mathbf{x}_i = \operatorname{vec}(\mathbf{X}_i) = \sum_\ell \mathbf{e}_\ell \otimes \mathbf{x}_{i\ell}. \tag{2}$$

We denote the centering matrix of order T by $\mathbf{A} = \mathbf{I}_T - \boldsymbol{\iota}_T\boldsymbol{\iota}'_T/T$ and write $\boldsymbol{\Delta}_j = \mathbf{D}_j\mathbf{D}'_j$, with $\mathbf{D}_j$ the $T \times (T - j)$ matrix that takes differences over time span $j = 1, \dots, T - 1$.

The within estimator is the MM estimator of $\boldsymbol{\beta}$ corresponding to the k 'within' moment conditions

$$\mathbb{E}(\mathbf{X}'_i\mathbf{A}\boldsymbol{\varepsilon}_i) = \mathbf{0}. \tag{3}$$

Evidently, we cannot use the J-test for an exactly identified system of moment conditions. One option would be to test the $T(T-1)k$ population moment conditions for exogeneity of $\mathbf{X}_i$, after removal of the individual effects by taking first differences. The resulting 'generic' moment conditions for exogeneity are given by

$$\mathbb{E}(\mathbf{x}_i \otimes \mathbf{D}'_1\boldsymbol{\varepsilon}_i) = \mathbf{0}. \tag{4}$$

These are not the moment conditions that we will use, though. We propose a J-test for the $(T - 1)k$

‘differences’ moment conditions given by

$$\mathbb{E}(\mathbf{X}_i'\boldsymbol{\Delta}_j\boldsymbol{\varepsilon}_i) = \mathbf{0} \qquad [j = 1, \dots, T-1]. \tag{5}$$

Before we turn to the rationale for our choice of moment conditions, we need to understand the relation among the aforementioned moment conditions. We start by noting that $\mathbf{A} = \mathbf{D}_1\boldsymbol{\Delta}_1^{-1}\mathbf{D}_1'$, because both $\mathbf{A}$ an $\mathbf{D}_1\boldsymbol{\Delta}_1^{-1}\mathbf{D}_1'$ are symmetric, idempotent of rank $T-1$ and orthogonal to $\boldsymbol{\iota}_T$. Consequently, the generic moment conditions in (4) are equivalent to

$$\mathbb{E}(\mathbf{x}_i \otimes \mathbf{A}\boldsymbol{\varepsilon}_i) = \mathbf{0}. \tag{6}$$

The within conditions in (3) are a linear combination of the conditions in (6). To see this, note that

$$(\mathbf{I}_k \otimes \mathrm{vec}(\mathbf{I}_T))'(\mathbf{x}_i \otimes \mathbf{A}\boldsymbol{\varepsilon}_i) = \mathbf{X}_i'\mathbf{A}\boldsymbol{\varepsilon}_i. \tag{7}$$

The generic moment conditions in (4) imply that

$$\mathbb{E}(\mathbf{x}_i \otimes \mathbf{D}_j'\boldsymbol{\varepsilon}_i) = \mathbf{0} \qquad [j = 1, \dots, T-1], \tag{8}$$

since the conditions in (8) are a linear combination of those in (4). That is, each $\mathbf{D}_j$ is a linear combination of the columns of $\mathbf{D}_1$: $\mathbf{D}_2$ is obtained by adding up each set of two adjacent columns of $\mathbf{D}_1$, $\mathbf{D}_3$ by adding up each set of three adjacent columns of $\mathbf{D}_1$, and so on. By noting that (8) is equivalent with $\mathbb{E}(\mathbf{x}_i \otimes \boldsymbol{\Delta}_j\boldsymbol{\varepsilon}_i) = \mathbf{0}$ we obtain, analogously to (7), the differences moment conditions in (5).

Because $\boldsymbol{\Delta}_j$ has j-th pseudo-diagonal equal to -1, it follows that $\mathbf{A} = (\boldsymbol{\Delta}_1 + \ldots + \boldsymbol{\Delta}_{T-1})/T$. All other pseudo-diagonals are zero, while $\sum_j \boldsymbol{\Delta}_j$ has diagonal elements equal to $T-1$ since all rows add to zero. Hence, the within conditions in (3) are also a linear combination of the differences moment conditions in (5). In particular, (3) is the average of (5) over j.

In sum, we have (4) $\equiv$ (6) $\Longrightarrow$ (8) $\Longrightarrow$ (5) $\Longrightarrow$ (3) due to the linear dependence of the various moment conditions.

2.2. Optimality

We consider the J-test for the differences moment conditions in (5), based on the $(1-\alpha)$-critical value of the central chi-square distribution with $(T-1)k$ degrees of freedom, for $0 < \alpha < 1$. For the sake of comparison, we also consider the J-test for the generic moment conditions in (4) based on the $(1-\alpha)$-critical value of the central chi-square distribution with $T(T-1)k$ degrees of freedom.

We are interested in the properties of both J-tests as a test of the within moment conditions in (3). In terms of asymptotic size, both tests are conservative, because they both test a set of moment conditions that imply the within moment conditions. That is, the rejection rates will be above nominal if (3) holds. Furthermore, the J-test for the generic moment conditions in (4) is more conservative than the J-test for the differences conditions (5), because (4) $\Longrightarrow$ (5).

Comparing both tests in terms of local power requires some more work. Because the J-test for the generic moment conditions in (4) has the maximum number of degrees of freedom, it has the largest possible value of the non-centrality parameter for all local alternatives (Newey, 1985, Prop. 6). Stated differently, for each local alternative, its non-centrality parameter is larger than or equal to the non-centrality parameter of the J-test for the differences moment conditions in (5). Hence, viewed as a test of the generic moment conditions, neither of the two tests has uniformly higher local power than the other one. This follows because the value of a given tail probability of the non-central chi-square distribution decreases with the number of degrees of freedom, but increases with the non-centrality parameter (Newey, 1985, p. 238).

However, we are interested in the local power of both tests if we view them as a test of the within conditions in (3). The J-test for the differences moment conditions in (5) turns out to be the more powerful J-test for local alternatives such that the within conditions in (3) do not hold. Intuitively, this J-test directs the power from the full set of moment conditions for exogeneity to the ones that are relevant for the within estimator. That is, if (3) does not hold locally, testing more moment conditions than those in (5) will not lead to more power. Furthermore, if (3) holds, testing more moment conditions than those in (5) will increase the risk of a rejection.

The theorem below provides a formal comparison of the two J-tests' local power properties as a test of (3), while assuming that both tests use the same $\alpha > 0$. The proof can be found in Section A of the appendix with supplementary material.

Theorem 1 (comparison of J-tests)
(i) For all local alternatives such that within moment conditions in (3) do not hold, the J-test for the differences moment conditions in (5) has higher power than the J-test for the generic moment conditions in (4). (ii) Both tests are conservative for local alternatives such that the within moment conditions in (3) hold, with local power larger than or equal to α.

The theorem tells us that, in order to detect local alternatives for which the within estimator's moment conditions do not hold, it is optimal to use the J-test for the differences conditions in (5) instead of the generic conditions in (4).

2.3. Wald test

The test we have proposed is a J-test for the differences moment conditions in (5), but we will use this test in the representation of a (numerically identical) Wald test. This has certain advantages, as we will explain below.

The idea of the Wald test is that we estimate $\boldsymbol{\beta}$ separately for each time span $j = 1, \ldots, T-1$ and then assess whether the resulting $T-1$ estimates $\hat{\boldsymbol{\beta}}_j$ are identical. With $\boldsymbol{\beta}_j = \operatorname{plim}_{n\to\infty} \hat{\boldsymbol{\beta}}_j$, our null hypothesis becomes $H_0 : \boldsymbol{\beta}_j = \boldsymbol{\beta}_{j+1}$, while the alternative hypothesis is $H_1 : \boldsymbol{\beta}_j \neq \boldsymbol{\beta}_{j+1}$ for at least one j ($j = 1, \ldots, T-2$).

Because the Wald test requires the joint asymptotic covariance matrix of the $\hat{\boldsymbol{\beta}}_j$s, we obtain the $\hat{\boldsymbol{\beta}}_j$s jointly as an exactly identified (G)MM estimator and use a cluster-robust estimator for the joint asymptotic covariance matrix. This yields the estimators

$$\hat{\boldsymbol{\beta}}_j = \left(\sum_i \mathbf{X}_i' \boldsymbol{\Delta}_j \mathbf{X}_i \right)^{-1} \sum_i \mathbf{X}_i' \boldsymbol{\Delta}_j \mathbf{y}_i, \tag{9}$$

for $j = 1, \ldots, T-1$. We store the $\hat{\boldsymbol{\beta}}_j$s in a $(T-1)k$-vector and denote this vector by $\hat{\boldsymbol{\beta}} = (\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_{T-1}')'$. Let $\mathbf{B}_1$ be the $(T-1) \times (T-2)$ matrix taking first differences and let $\mathbf{R} = \mathbf{B}_1 \otimes \mathbf{I}_k$. The Wald statistic corresponding to H_0 is given by

$$q_{\mathrm{w}} = \hat{\boldsymbol{\beta}}' \mathbf{R} \left(\sum_i \mathbf{R}' \mathbf{u}_i \mathbf{u}_i' \mathbf{R} \right)^{-1} \mathbf{R}' \hat{\boldsymbol{\beta}}, \tag{10}$$

where

$$\mathbf{u}_i = \begin{pmatrix} \left(\sum_\ell \mathbf{X}_\ell' \boldsymbol{\Delta}_1 \mathbf{X}_\ell\right)^{-1} \mathbf{X}_i' \boldsymbol{\Delta}_1 \hat{\boldsymbol{\varepsilon}}_{i1} \\ \vdots \\ \left(\sum_\ell \mathbf{X}_\ell' \boldsymbol{\Delta}_{T-1} \mathbf{X}_\ell\right)^{-1} \mathbf{X}_i' \boldsymbol{\Delta}_{T-1} \hat{\boldsymbol{\varepsilon}}_{i,T-1} \end{pmatrix}, \tag{11}$$

with $\hat{\boldsymbol{\varepsilon}}_{ij} = \mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\beta}}_j$ for $j = 1, \ldots, T-1$.

Because of the linearity of the moment conditions in $\boldsymbol{\beta}$, q_{w} is numerically identical to the J-test statistic for the differences conditions in (5), provided that both statistics use the same consistent estimator for the covariance matrix. This equality holds both under the null and the (fixed or local) alternative hypothesis. Furthermore, the Wald test statistic is also identical to three other well-known tests for $H_1 : \boldsymbol{\beta}_j = \boldsymbol{\beta}_{j+1}$ ($j = 1, \ldots, T-2$): the LM test, the distance-difference test (i.e., the GMM equivalent of the likelihood-ratio test) and the minimum chi-square test. Again the equality only holds if the same consistent estimator for the covariance matrix is used (Newey and West, 1987; Newey and

McFadden, 1994; Ruud, 2000).

Because the above tests are numerically identical, the choice for any one representation can be made on non-statistical grounds (Newey and West, 1987). We choose the Wald test because of its appealing economic interpretation and its potential for visualization. By using the Wald test, the violation of overidentifying restrictions is translated into patterns in the $\boldsymbol{\beta}_j$s. If H_0 is rejected, the patterns in the $\hat{\boldsymbol{\beta}}_j$s reveal the economic relevance of the within estimator's inconsistency. If the variation in the $\hat{\boldsymbol{\beta}}_j$s is small and the $\hat{\boldsymbol{\beta}}_j$s are close to the within estimator, the economic importance of the rejection will be limited. An informal visualization of the Wald test is obtained by plotting $\hat{\boldsymbol{\beta}}_j$s as a function of j for each covariate, with the within estimator added as a horizontal line. We will refer to this as the 'difference curves'.

2.4. Motivating examples

To illustrate the link between the Wald test and familiar situations where the within estimator is inconsistent, Table 1 considers four cases: (i) classical measurement error, (ii) non-classical measurement error, (iii) omitted variables and (iv) simultaneity. The precise model specification in each case is described in the first column of Table 1. We emphasize that the list of examples is not exhaustive; evidently, there are many other cases in which the within estimator is inconsistent. All calculations related to Table 1 have been relegated to Sections B and C of the appendix with supplementary material.

In each case, a univariate version of the linear panel regression model in (1) is estimated using the within and differences estimators. The resulting estimators of the regression coefficient are invariably inconsistent. The second column in Table 1 reports the inconsistency of the differences estimator in each case.

The last two columns pertain to the local power of the Wald test for $T = 3$. We first mention the local alternatives with Pitman drift considered in each case. The final column shows the implied non-centrality parameter of the Wald test under these local alternatives.[1]

The local power of the Wald test arises from the differences in the probability limits β_j for different values of j. For certain parameter values, however, the inconsistencies do not depend on j. An example of such a case is classical measurement error with equal persistence of the unobserved regressor and the measurement error. The non-centrality parameter is 0 in such cases, while the local power is equal to α (also referred to as 'trivial' power). Because of the equality of the Wald and J-tests, the cases where the Wald test has trivial power correspond to the cases where also the J-test has trivial

[1]Section C of the supplementary material provides expressions for the non-centrality parameter for $T \geq 4$, which turn out similar as for $T = 3$.

power. We refer to Newey (1985, Prop. 1) and Hall (2005, Th. 5.4) for more details about the local alternatives under which the J-test has trivial power.

We have performed a simulation study for each of the motivating examples to investigate the finite-sample properties of the Wald test in terms of power and size under fixed alternatives. The simulation results confirm the Wald tests's consistency under fixed alternatives: even a relatively weak pattern in the $\boldsymbol{\beta}_j$s can be detected, provided that n is large enough. These results can be found in Section D of the supplementary material.

2.5. *Empirical strategy*

Turning back to the general case, we propose an empirical strategy based on the proposed Wald test:

(1) Estimate q_{w}. If q_{w} exceeds the critical value of the $\chi^2_{(T-2)k}$ distribution, then reject $H_0 : \boldsymbol{\beta}_j = \boldsymbol{\beta}_{j+1}$ $(j = 1, \dots, T-2)$.

(2) If H_0 is rejected, plot the difference curves for all covariates and verify the economic significance of the non-constant patterns in the difference curves, especially if n is large.

(3) In case of both statistical and economic significance, revise the set of moment conditions and test again. For this purpose, assume that there are m potentially endogenous covariates, for which there are m candidate instrumental variables available. Then return to Step 1, but replace the differences moment conditions for the m covariates by the differences moment conditions for the m instruments.

(4) If there are more than m candidate instruments available for the m potentially endogenous covariates, then use a standard J-test for the within moment conditions. Because of the overidentification, it is no longer necessary to use the larger set of differences moment conditions.

(5) If no candidate instruments are available or if H_0 continues to be rejected, then switch to an estimator that requires less stringent exogeneity assumptions than the within estimator, or change the model specification using any available information about the nature of the endogeneity.

Steps 1–2 apply to running and visualizing our test and do not require any prior information about the nature of the endogeneity. Evidently, if the model has to be revised because H_0 is rejected, the use of instrumental variables will require some idea about the potentially endogenous variables.

3. Empirical application

The consistent estimation of cost and production functions is a long-standing problem in econometrics (e.g., Coen and Hickman, 1970; McElroy, 1987; Griliches and Hausman, 1986; Mundlak, 1996; Paris and Caputo, 2004; Dimitropoulos, 2015). In particular, Griliches and Hausman (1986) investigate this problem in the context of the 'short run increasing returns to scale puzzle' for manufacturing firms' labor elasticity. This puzzle refers to estimated labor elasticities of output that are less than unity, which is economically implausible because it would indicate increasing returns to scale to labor alone. The empirical application that we provide in this section fits in this strand of literature. We will estimate U.S. banks' cost functions and the implied scale elasticities using the within estimator and use our test to investigate the estimator's consistency.

3.1. Banks' scale elasticity

Banks' scale effects are typically measured by the scale elasticity. This elasticity is the inverse of the cost elasticity with respect to output (Hanoch, 1975). A scale elasticity larger than one indicates the presence of economies of scale, meaning that banks' unit costs of production decrease with output. Many recent banking studies provide estimates of banks' scale elasticities and tend to find scale elasticity estimates that are significantly larger than one (e.g., Feng and Serletis, 2010; Wheelock and Wilson, 2012; Hughes and Mester, 2013; Feng and Zhang, 2012, 2014; Beccalli et al., 2015; Spierdijk and Zaouras, 2018; Wheelock and Wilson, 2018). Scale elasticities play an important role in banking and are considered to have a high policy relevance. For example, Hughes and Mester (2013) plead against a size limit on banks on the basis of their scale elasticity estimates, while Boyd and Heitz (2016) use scale elasticity estimates to determine the implied net costs of increasing bank size for too-big-to-fail banks.

3.1.1. Cost function

We follow the intermediation model of banking (Klein, 1971; Monti, 1972; Sealey and Lindley, 1977) and assume that banks employ a cost technology with three input factors (funding, personnel, and physical capital) and total assets as the single output factor (q_{it}). For bank $i = 1, \ldots, n$ in year $t = 1, \ldots, T$, the corresponding input-factor prices are the price of funding ($p_{1,it}$), the wage rate ($p_{2,it}$), and the price of physical capital ($p_{3,it}$). Total input-factor costs (c_{it}) are defined as the sum of expenses on funding, personnel, and physical capital. The quantity of total assets is denoted by q_{it}.

We model the dependence of total input-factor costs on input-factor prices and total assets using a translog cost function, which was introduced by Christensen et al. (1971, 1973). This type of cost function provides a log-quadratic approximation to a true cost function and has been widely used virtually all areas of economics ever since (e.g., Koetter et al., 2012; Byrne, 2015; Grieco et al., 2016; Kee and Tang, 2016; Krasnokutskaya et al., 2018).

We initially consider a simplified version of the quadratic translog cost function for sake of exposition, which results in a multivariate linear panel regression model. More specifically, we consider the following three-input and one-output translog cost function for bank i in year t:

$$\log(\widetilde{c}_{it}) = \alpha_i + \gamma_t + \beta_q \log(q_{it}) + \frac{1}{2}\beta_{qq}[\log(q_{it})]^2 + \sum_{k=2}^{3} \beta_{pk} \log(\widetilde{p}_{k,nt}) + \varepsilon_{it}, \tag{12}$$

where α_i denotes a bank-specific effect that is potentially correlated with the error term ε_{it} and γ_t a year fixed effect. Throughout, variables with a tilde have been divided by the price of funding $p_{1,it}$ to ensure that the cost function features linear homogeneity in input prices.

The implied scale elasticity for the simple translog cost function in (12) equals

$$e(q_{it}) = \left(\frac{\partial \log(c_{it})}{\partial \log(q_{it})}\right)^{-1} = \frac{1}{\beta_q + \beta_{qq}\log(q_{it})}. \tag{13}$$

In the usual case that average costs are U-shaped (i.e., $\beta_{qq} > 0$), the scale elasticity is a decreasing function of (log) output. Throughout, we will evaluate the scale elasticity in the sample mean of log output and denote the resulting scale elasticity estimate by $\bar{e}$.

3.1.2. Potential sources of endogeneity

There are several reasons to believe that the within estimator of (12) is inconsistent. Although cost functions in terms of observed input prices and outputs are still widely used in the literature, it has been known for long that this approach is problematic from the perspective of measurement error (e.g., Coen and Hickman, 1970; McElroy, 1987; Griliches and Hausman, 1986; Mundlak, 1996; Paris and Caputo, 2004; Dimitropoulos, 2015). In reality, the demand for input factors will be based on expected output levels and input prices. Consequently, using observed values instead of expected values in the cost function will result in measurement error. Furthermore, the effect of measurement error in the output variable will be exacerbated in the presence of a quadratic term (Griliches and Ringstad, 1970).

There are also two potential sources of omitted variables. The first source relates to functional misspecification of the translog cost function. A full quadratic translog cost function contains more terms than the ones included in (12), which we omitted for the sake of exposition. Even third- or higher-order terms may be required to provide an accurate fit to the data. These and other forms of functional misspecification of the regression function can be viewed as a form of omitted variables bias (Plosser et al., 1982). The second source of omitted variables relates to bank-specific control variables. For example, the simple translog cost function in (12) does not control for bank risk, asset quality and other time-varying bank characteristics (Mester, 1996).

3.1.3. Estimation results

We use year-end 2011–2017 Call Report data, consisting of banks' balance sheets and income statements. We construct a balanced annual sample of $n = 2,505$ U.S. banks covering $T = 7$ years, with a total of 17,535 bank-year observations. Section E of the supplementary material explains the selection of banks in more detail. Using the series available in the Call Report data, we calculate input prices in a way that is common in banking. This is also explained in the appendix.

We include time fixed effects in all specifications and estimate the translog cost function using the within and differences estimators. The coefficient estimates and the implied estimates of $\bar{e}$ are reported in Table 2.

Our Wald statistic has a value of 374.5. With a critical value of 31.4 (degrees of freedom 20), the null hypothesis $H_0 : \beta_{q,j} = \beta_{q,j+1}; \beta_{qq,j} = \beta_{qq,j+1}; \beta_{p2,j} = \beta_{p2,j+1}; \beta_{p3,j} = \beta_{p3,j+1}$ $(j = 1, \ldots, T-2)$ is rejected at each reasonable significance level.[2]

Figure 1 shows the difference curve for the scale elasticity $\bar{e}$. This elasticity is our main object of interest and a function of the two coefficients related to output; see (13). The scale elasticity implied by the within estimator equals 1.23 and is significantly larger than one, suggesting substantive scale economies. The j-th differences estimates of $\bar{e}$ decrease with j from 1.28 to 1.16, confirming the economic relevance of the statistical rejection. Because our Wald test provides strong evidence against the consistency of the within estimator, we cannot rely on the implied estimate of the scale elasticity.

3.1.4. Robustness checks

For the sake of exposition, we have used a simple translog cost function. We have also estimated a wide range of additional specifications, including: (1) different samples, (2) more complete quadratic translog costs functions, (3) the inclusion of time-varying bank-specific control variables such as the equity ratio, (4) multiple outputs (such as loans, securities and off-balance sheet activities) instead of total assets as the single aggregate output, (5) stratified estimation on the basis of total bank output to account for differences in cost technology between banks of different sizes, and (6) an alternative functional form known as the generalized Leontief cost function (Diewert, 1971; Hall, 1973; Diewert, 1976; Fuss, 1977). Leontief technologies have been widely used in banking research and other fields (e.g., Thomsen, 2000; Gunning and Sickles, 2011; Martín-Oliver et al., 2013; Miller et al., 2013).

[2] Because the Wald test is numerically identical to a J-test, we may sometimes need a generalized inverse to calculate the statistic (Newey, 1985). We note that the J-test statistic does not depend on the choice of generalized inverse due to the linearity of our moment conditions (Newey, 1985).

All these additional specifications are still linear panel regression models to which our test can be applied. In all cases, we continue to find strong evidence against the within estimator's consistency. A selection of the additional estimation results can be found in Section F of the supplementary material.

3.1.5. Consistent estimation

We follow the empirical strategy outlined in Section 2.5 and consider potential instruments for the covariates related to input prices and output. Finding such instruments turns out challenging. Although a bank's amount of fixed assets and its total number of full-time equivalent employees seem candidate supply-side instruments for output, both of them are already used in the model. Labor is considered to be an input factor, while the number of full-time equivalent employees is used to calculate the wage rate. Furthermore, fixed assets are contained in total assets. Demand-side candidate instruments for output, including measures of economic activity and the yield on e.g. Treasury bills, do not vary across banks and turn out weak instruments. Similarly, it is challenging to find instruments for input prices. As an alternative, we could resort to modeling banks' cost-minimizing input decisions on the basis of their expectations concerning input prices and output. This would yield a system of non-linear equations with latent variables and measurement error (e.g., Paris and Caputo, 2004; Dimitropoulos, 2015).

For any alternative approach, the limitations of the translog cost function to proxy the shape of an unknown cost function should be a point of attention (e.g., White, 1980; Gallant, 1981; Byron and Bera, 1983; Bera, 1984; Aizcorbe, 1992; McAllister and McManus, 1993). Because of these limitations, several studies have used non-parametric techniques to estimate the cost function and the implied scale elasticity (e.g., Wheelock and Wilson, 2012, 2018). Because measurement error in observed output quantities and input prices arises independently of the functional form of the cost function, it remains important to deal with measurement error even if the cost function is estimated non-parametrically. More specifically, Driscoll and Boisvert (1991) show that more complex functional forms could actually do more harm than good in the presence of ignored measurement error. They apply both second- and third-order translog functions to simulated data with and without measurement error. They show that the third-order models do not outperform the quadratic models in the presence of measurement error, while they do in the error-free case.

All in all, these considerations show that 'fixing' the cost model is complicated and suggest that it is by no means guaranteed that a correctly specified model will eventually be found. Given the substantial societal relevance of banks' scale economies, it nevertheless remains important to continue the quest for a correct specification.

4. Conclusion

The within estimator is widely used to estimate the linear panel regression model. Its consistency requires exogeneity of the regressors, which can be violated due to e.g. measurement error, omitted variables and simultaneity.

We have exploited the differences between the short- and long-differences estimators in the linear panel data model to construct a GMM-test for the consistency of the within estimator. This test is locally more powerful than a more generic GMM-test for exogeneity of the regressors. Throughout, we have used our GMM-test in the representation of a Wald test, which facilitates the economic interpretation and visualization of the test outcomes using 'difference curves'.

If our test fails to reject, there is no evidence against the within estimator's consistency. Although this is the most favorable outcome, researchers should be aware of the possibility that the test may have low power in certain cases. It therefore remains important to look for other evidence against the within estimator, such as coefficient signs and magnitudes that are unlikely from an economic perspective. Researchers should also be aware of the possibility that low power could arise from limited data variability due to taking differences, yielding coefficient estimates with relatively large standard errors.

If our test rejects, we recommend several salvaging steps to achieve consistent estimation. As usual, however, finding a well-specified model remains to a large extent a case-by-case puzzle without guaranteed success, depending on e.g. prior information and the availability of valid and strong instruments. What is universal, though, is our urgent advice to researchers working with panel data to routinely run our test and to draw the associated difference curves, and to discard the within estimator for further inference if it fails to pass the test.

References

Aizcorbe, A., 1992. A note on empirical tests of separability and the "approximation" view of functional forms. Southern Economic Journal 59, 310–318.

Amini, S., Delgado, M., Henderson, D., Parmeter, C., 2012. Fixed vs random: The Hausman test four decades later, in: Baltagi, B., Carter Hill, R., Newey, W., White, H. (Eds.), Essays in Honor of Jerry Hausman (Advances in Econometrics, Vol. 29). Emerald Group Publishing Limited, pp. 479–513.

Arellano, M., Bond, S., 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. Review of Economic Studies 58, 277–297.

Baltagi, B., Bresson, G., Pirotte, A., 2003. Fixed effects, random effects or Hausman-Taylor? A pretest estimator. Economics Letters 79, 361 – 369.

Beccalli, E., Anolli, M., Borello, G., 2015. Are European banks too big? Evidence on economies of scale. Journal of Banking & Finance 58, 232–246.

Bera, A., 1984. The use of linear approximation to nonlinear regression analysis. Sankhyā: The Indian Journal of Statistics, Series B 46, 285–290.

Boyd, J., Heitz, A., 2016. The social costs and benefits of too-big-to-fail banks: A "bounding" exercise. Journal of Banking & Finance 68, 251–265.

Breusch, T., Godfrey, L., 1986. Data transformation tests. Economic Journal 96, 47–58.

Bun, M., Kelaher, R., Sarafidis, V., Weatherburn, D., 2019. Crime, deterrence and punishment revisited. Empirical Economics forthcoming.

Byrne, D., 2015. Testing models of differentiated products markets: Consolidation in the cable T.V. industry. International Economic Review 56, 805–850.

Byron, R., Bera, A., 1983. Least squares approximations to unknown regression functions: A comment. International Economic Review 24, 255–260.

Christensen, L., Jorgenson, D., Lau, L., 1971. Conjugate duality and the transcendental logarithmic function. Econometrica 39, 255–256.

Christensen, L., Jorgenson, D., Lau, L., 1973. Transcendental logarithmic production frontiers. Review of Economics and Statistics 55, 28–45.

Coen, R., Hickman, B., 1970. Constrained joint estimation of factor demand and production functions. Review of Economics and Statistics 52, 287–300.

Davidson, R., Godfrey, L., MacKinnon, J., 1985. A simplified version of the differencing test. International Economic Review 26, 639–647.

Diewert, W., 1971. An application of the Shephard duality theorem: A generalized Leontief production function. Journal of Political Economy 79, 481–507.

Diewert, W., 1976. Exact and superlative index numbers. Journal of Econometrics 4, 115–145.

Dimitropoulos, D., 2015. Three Essays in Energy Economics and Industrial Organization, with Applications to Electricity and Distribution Networks. PhD Thesis, Department of Economics, University of Toronto.

Driscoll, P., Boisvert, R., 1991. Dual second- and third-order translog models of production. American Journal of Agricultural Economics 73, 1146–1160.

Feng, G., Serletis, A., 2010. Efficiency, technical change, and returns to scale in large U.S. banks: Panel data evidence from an output distance function satisfying theoretical regularity. Journal of Banking & Finance 34, 127–138.

Feng, G., Zhang, X., 2012. Productivity and efficiency at large and community banks in the U.S.: A Bayesian true random effects stochastic distance frontier analysis. Journal of Banking & Finance 36, 1883–1895.

Feng, G., Zhang, X., 2014. Returns to scale at large banks in the US: A random coefficient stochastic frontier approach. Journal of Banking & Finance 39, 135–145.

Fuss, M., 1977. The structure of technology over time: A model for testing the "putty-clay" hypothesis. Econometrica 45, 1797–1821.

Gallant, A., 1981. On the bias in flexible functional forms and an essentially unbiased form: The Fourier flexible form. Journal of Econometrics 15, 211–245.

Goolsbee, A., 2000. The importance of measurement error in the cost of capital. National Tax Journal 53, 215–228.

Grieco, P.E., Li, S., Zhang, H., 2016. Production function estimation with unobserved input price dispersion. International Economic Review 57, 665–690.

Griliches, Z., Hausman, J., 1986. Errors in variables in panel data. Journal of Econometrics 31, 93–118.

Griliches, Z., Ringstad, V., 1970. Error-in-the-variables bias in nonlinear contexts. Econometrica 38, 368–370.

Gunning, T., Sickles, R., 2011. A multi-product cost function for physician private practices. Journal of Productivity Analysis 35, 119–128.

Hall, A., 2005. Generalized method of moments. Oxford University Press.

Hall, R., 1973. The specification of technology with several kinds of output. Journal of Political Economy 81, 878–892.

Hanoch, G., 1975. The elasticity of scale and the shape of average costs. American Economic Review 65, 492–497.

Harris, R., Tzavalis, E., 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics 91, 201 – 226.

Hausman, J., 1978. Specification tests in econometrics. Econometrica 46, 1251–1271.

Hayakawa, K., 2019. Alternative over-identifying restriction test in the GMM estimation of panel data models. Econometrics and Statistics 10, 71 – 95.

Hughes, J., Mester, L., 2013. Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. Journal of Financial Intermediation 22, 559–585.

Joshi, R., Wooldridge, J., 2019. Correlated random effects models with endogenous explanatory variables and unbalanced panels. Annals of Economics and Statistics , 243–268.

Kee, H., Tang, H., 2016. Domestic value added in exports: Theory and firm evidence from China. American Economic Review 106, 1402–1436.

Klein, M., 1971. A theory of the banking firm. Journal of Money, Credit and Banking 3, 205–218.

Koetter, M., Kolari, J., Spierdijk, L., 2012. Enjoying the quiet life under deregulation? Evidence from adjusted Lerner indices for US banks. Review of Economics and Statistics 94, 462–480.

Krasnokutskaya, E., Li, Y., Todd, P., 2018. Product choice under government regulation: The case of chile's privatized pension system. International Economic Review 59, 1747–1783.

Levitt, S., 1998. Why do increased arrest rates appear to reduce crime: Deterrence, incapacitation, or measurement error? Economic Inquiry 36, 353–372.

Martín-Oliver, A., Ruano, S., Salas-Fumás, V., 2013. Why high productivity growth of banks preceded the financial crisis. Journal of Financial Intermediation 22, 688–712.

McAllister, P., McManus, D., 1993. Resolving the scale efficiency puzzle in banking. Journal of Banking & Finance 17, 389–405.

McElroy, M., 1987. Additive general error models for production, cost, and derived demand or share systems. Journal of Political Economy 95, 737–757.

McKinnish, T., 2008. Panel data models and transitory fluctuations in the explanatory variable, in: Fomby, T., Hill, R., Millimet, D., Smith, J., Vytlacil, E. (Eds.), Advances in Econometrics. volume 21: Modelling and Evaluating Treatment Effects in Econometrics, pp. 335–358.

Mester, L., 1996. A study of bank efficiency taking into account risk-preferences. Journal of Banking & Finance 20, 1025–1045.

Miller, M., Zhang, L., Li, H., 2013. When bigger isn't better: Bailouts and bank reform. Oxford Economic Papers 65, i7–i41.

Monti, M., 1972. Deposit, credit, and interest rate determination under alternative bank objectives, in: Szegö, G., Shell, K. (Eds.), Mathematical Methods in Investment and Finance. North-Holland, pp. 431–454.

Mundlak, Y., 1961. Empirical production function free of management bias. Journal of Farm Economics 43, 44–56.

Mundlak, Y., 1996. Production function estimation: Reviving the primal. Econometrica 64, 431–438.

Mundlak, Y., Hoch, I., 1965. Consequences of alternative specifications in estimation of Cobb-Douglas production functions. Econometrica 33, 814–828.

Newey, W., 1985. Generalized method of moments specification testing. Journal of Econometrics 29, 229–256.

Newey, W., McFadden, D., 1994. Large sample estimation and hypthesis testing, in: Griliches, Z., Intriligator, M. (Eds.), Handbook of Econometrics, Volume 4. North Holland.

Newey, W., West, K., 1987. Hypothesis testing with efficient method of moments estimation. international Economic Review 28, 777–787.

Paris, Q., Caputo, M., 2004. Efficient estimates of a model of production and cost URL:

https://ageconsearch.umn.edu/record/93742/files/04-016.pdf.

Plosser, C., Schwert, G., White, H., 1982. Differencing as a test of specification. International Economic Review 23, 535–552.

Ruud, P., 2000. An introduction to classical econometric theory. Oxford University Press.

Sarafidis, V., Wansbeek, T., 2012. Cross-sectional dependence in panel data analysis. Econometric Reviews 31, 483–531.

Sealey, C., Lindley, J., 1977. Inputs, outputs, and a theory of production and cost at depository financial institutions. Journal of Finance 32, 1251–1266.

Spierdijk, L., Zaouras, M., 2018. Measuring banks market power in the presence of economies of scale: A scale-corrected Lerner index. Journal of Banking & Finance 87, 40–48.

Thomsen, T., 2000. Short cuts to dynamic factor demand modelling. Journal of Econometrics 97, 1–23.

Thursby, J., 1989. A comparison of several specification error tests for a general alternative. International Economic Review 30, 217–230.

Verbeek, M., Nijman, T., 1992. Testing for selectivity bias in panel data models. International Economic Review 33, 681–703.

Wheelock, D., Wilson, P., 2012. Do large banks have lower costs? New estimates of returns to scale for U.S. banks. Journal of Money, Credit and Banking 44, 171–199.

Wheelock, D., Wilson, P., 2018. The evolution of scale economies in U.S. banking. Journal of Applied Econometrics 33, 16–28.

White, H., 1980. Using least squares to approximate unknown regression functions. International Economic Review 21, 149–170.

Wooldridge, J.M., 1995. Selection corrections for panel data models under conditional mean independence assumptions. Journal of Econometrics 68, 115–132.

Table 1: Motivating examples: inconsistency, local alternative and non-centrality parameter

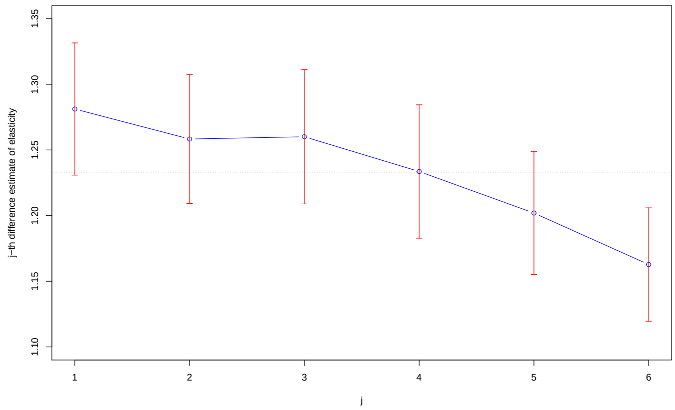
Model	Inconsistency	H_1	Non-centrality ($T=3$)
CLASSICAL ME $y_{it} = \alpha_i + \beta\xi_{it} + \varepsilon_{it}$ $x_{it} = \xi_{it} + v_{it}$ $\xi_{it} = \rho\xi_{i,t-1} + \theta_{it}$ $v_{it} = \delta v_{i,t-1} + \eta_{it}$ $\sigma_{\theta\eta} = 0$	$-\beta \dfrac{\sigma_\eta^2 \frac{1-\delta^j}{1-\delta^2}}{\sigma_\theta^2 \frac{1-\rho^j}{1-\rho^2} + \sigma_\eta^2 \frac{1-\delta^j}{1-\delta^2}}$	$\sigma_\eta^2 = \dfrac{\tilde{\sigma}_\eta^2}{\sqrt{n}}$	$\dfrac{\beta^2\tilde{\sigma}_\eta^4}{\sigma_\theta^2\sigma_\varepsilon^2} \dfrac{8(\delta-\rho)^2}{(1+\delta)^2(9-\rho^2)}$
NON-CLASSICAL ME as classical, but $\sigma_{\theta\eta} \neq 0$ $W_j \equiv \dfrac{1-(\delta^j+\rho^j)/2}{1-\delta\rho}$ $W \equiv \dfrac{1-\delta^2}{1-\delta\rho}$	$-\beta \dfrac{\sigma_\eta^2 \frac{1-\delta^j}{1-\delta^2} + \sigma_{\theta\eta} W_j}{\sigma_\theta^2 \frac{1-\rho^j}{1-\rho^2} + \sigma_\eta^2 \frac{1-\delta^j}{1-\delta^2} + 2\sigma_{\theta\eta} W_j}$	$\sigma_\eta^2 = \dfrac{\tilde{\sigma}_\eta^2}{\sqrt{n}}$ $\sigma_{\theta\eta} = \dfrac{\tilde{\sigma}_{\theta\eta}}{\sqrt{n}}$	$\dfrac{2\beta^2}{\sigma_\theta^2\sigma_\varepsilon^2} \dfrac{(2\tilde{\sigma}_\eta^2 + \tilde{\sigma}_{\theta\eta} W)^2(\delta-\rho)^2}{(1+\delta^2)(9-\rho^2)}$
OMITTED VARIABLES $y_{it} = \alpha_i + \beta x_{it} + \gamma z_{it} + \varepsilon_{it}$ $x_{it} = \rho x_{i,t-1} + \theta_{it}$ $z_{it} = \delta z_{i,t-1} + \eta_{it}$ $\pi \equiv \gamma\sigma_{\theta\eta}$	$\gamma \dfrac{\sigma_{\theta\eta} \frac{1-(\delta^j+\rho^j)/2}{1-\delta\rho}}{\sigma_\theta^2 \frac{1-\rho^j}{1-\rho^2}}$	$\gamma = \dfrac{\tilde{\gamma}}{\sqrt{n}}$	$\dfrac{\tilde{\gamma}^2\sigma_{\theta\eta}^2\sigma_\theta^2}{\sigma_\eta^4\sigma_\varepsilon^2} \dfrac{2(1-\delta)^2(\delta-\rho)^2}{(1-\delta\rho)^2(9-\rho^2)}$
SIMULTANEITY $y_{it} = \beta_i + \beta x_{it} + \varepsilon_{it}$ $x_{it} = \alpha_i + \alpha y_{it} + u_{it}$ $u_{it} = \rho u_{i,t-1} + \theta_{it}$	$\dfrac{(1-\alpha\beta)\alpha\sigma_\varepsilon^2}{\alpha^2\sigma_\varepsilon^2 + \sigma_\theta^2 \frac{1-\rho^j}{1-\rho^2}}$	$\alpha = \dfrac{\tilde{\alpha}}{\sqrt{n}}$	$\dfrac{\tilde{\alpha}^2\sigma_\varepsilon^2}{\sigma_\theta^2} \dfrac{8\rho^2}{9-\rho^2}$

Table 2: Estimation results for the translog cost function

	FE	D_1	D_2	D_3	D_4	D_5	D_6
$\log(q)$	0.69	0.26	0.56	0.64	0.73	0.83	0.89
	(0.00)	(0.12)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$100 \times \frac{1}{2}[\log(q)]^2$	0.95	4.03	1.85	1.22	0.64	-0.02	-0.21
	(0.43)	(0.00)	(0.16)	(0.36)	(0.62)	(0.99)	(0.84)
$\log(\widetilde{p}_2)$	0.76	0.69	0.75	0.77	0.78	0.78	0.78)
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\log(\widetilde{p}_3)$	0.07	0.08	0.07	0.06	0.06	0.06	0.06
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\bar{e}$	1.23	1.28	1.26	1.26	1.23	1.20	1.16
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
adj. R^2	0.85	0.83	0.84	0.85	0.85	0.86	0.86

Notes: This table shows the estimated coefficients of the translog cost function in (12) applied to the U.S. banking data. The column captioned 'FE' reports the estimation results for the within estimator, while the columns captioned 'D_j' contain the j-th differences estimates. The p-values associated with the estimated coefficients are in parentheses. The time fixed effects are not reported, but have been taken out by using the dependent variable and the regressors in deviations from their means per time period. The associated scale elasticity in (13), evaluated in the sample mean of $\log(q_{it})$, is also shown. The reported p-value for the scale elasticity corresponds to a one-sided t-test of the null hypothesis $H_0 : \bar{e} = 1$ (no scale effects) against the alternative hypothesis $H_0 : \bar{e} > 1$ (economies of scale). The standard error of $\bar{e}$ used in this test is based on the Delta-method. The reported adjusted R^2s apply to the models in terms of the transformed data.

Figure 1: Difference curve for the scale elasticity $\bar{e}$



Notes: This figure shows the difference curve for the scale elasticity in (13), evaluated in the sample mean of $\log(q_{it})$. The intervals in red show the 95% asymptotic confidence interval for each point estimate. The dashed line indicates the value of the within estimator.

Supplementary Material

Laura Spierdijk[a,*], Tom Wansbeek[a]

[a] *University of Groningen, Faculty of Economics and Business, Department of Economics, Econometrics and Finance, P.O. Box 800, 9700 AV Groningen, The Netherlands.*

Abstract

This document contains the appendix with supplementary material belonging to "Differencing as a Consistency Test for the Within Estimator".

*Corresponding author

Email addresses: `L.Spierdijk@rug.nl` (Laura Spierdijk), `T.J.Wansbeek@rug.nl` (Tom Wansbeek)

Contents

A. Proof of Theorem 1

A.1. Preliminaries

For the sake of convenience, we repeat some notation from the main text. We consider the linear panel regression model with T observations over time,

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i, \tag{A.1}$$

for $i = 1, \ldots, n$; $\mathbf{y}_i$ $(T \times 1)$ is the dependent variable, $\mathbf{X}_i$ $(T \times k)$ is the matrix of observed covariates, $\boldsymbol{\beta}$ $(k \times 1)$ a coefficient vector, and $\boldsymbol{\varepsilon}_i$ $(T \times 1)$ the error term containing an individual effect possibly correlated with $\mathbf{X}_i$. Let $\mathbf{e}_\ell$ $(k \times 1)$ be the ℓ-th unit vector and write

$$\mathbf{X}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{ik}) = \sum_\ell \mathbf{x}_{i\ell}\mathbf{e}'_\ell \qquad \text{such that} \qquad \mathbf{x}_i = \operatorname{vec}(\mathbf{X}_i) = \sum_\ell \mathbf{e}_\ell \otimes \mathbf{x}_{i\ell}. \tag{A.2}$$

The centering matrix of order T is denoted by $\mathbf{A} = \mathbf{I}_T - \boldsymbol{\iota}_T\boldsymbol{\iota}'_T/T$, with $\boldsymbol{\iota}_T$ a T-vector of ones. Furthermore, we write $\boldsymbol{\Delta}_j = \mathbf{D}_j\mathbf{D}'_j$, with $\mathbf{D}_j$ the $T \times (T - j)$ matrix that takes differences over time span $j = 1, \ldots, T - 1$.

The within estimator is the exactly identified (G)MM estimator of $\boldsymbol{\beta}$ corresponding to the k 'within' moment conditions

$$\mathbb{E}(\mathbf{X}'_i\mathbf{A}\boldsymbol{\varepsilon}_i) = \mathbf{0}. \tag{A.3}$$

The $T(T-1)k$ 'generic' moment conditions for strict exogeneity of $\mathbf{X}_i$ after removing the individual effect are given by

$$\mathbb{E}(\mathbf{x}_i \otimes \mathbf{D}'_1\boldsymbol{\varepsilon}_i) = \mathbf{0}. \tag{A.4}$$

We propose a J-test for the $(T-1)k$ 'differences' moment conditions given by

$$\mathbb{E}(\mathbf{X}'_i\boldsymbol{\Delta}_j\boldsymbol{\varepsilon}_i) = \mathbf{0} \qquad [j = 1, \ldots, T-1]. \tag{A.5}$$

A.2. The proof

Let $p = T(T-1)k$ and $q = (T-1)k$. We consider the J-test for the differences moment conditions in (A.5). The GMM estimator of $\boldsymbol{\beta}$ in (A.1) that is used for this test is the two-step efficient GMM estimator corresponding to the differences conditions in (A.5), denoted $\hat{\boldsymbol{\beta}}_1$. The resulting J-test is referred to as the J_1-test. We assume that the J_1 test uses the $(1-\alpha)$ quantile of the chi-square distribution with $q-k$ degrees of freedom, for $0<\alpha<1$.

We also consider the J-test for the p generic moment conditions in (A.4). The GMM estimator used for this test is the two-step efficient GMM estimator of $\boldsymbol{\beta}$ in (A.1) corresponding to the generic moment conditions in (A.4). We denote the resulting GMM estimator by $\hat{\boldsymbol{\beta}}_2$. The resulting J-test is referred to as the J_2-test. The use of $\hat{\boldsymbol{\beta}}_1$ instead of $\hat{\boldsymbol{\beta}}_2$ for this J-test yields an asymptotically equivalent J-test, with the same non-centrality parameter as the initial J_2-test (Newey, 1983, Lemma 1.9). Because of this asymptotic equivalence, we will continue to refer to this test as the J_2-test. We assume that the J_2 test uses the $(1-\alpha)$ quantile of the chi-square distribution with $p-k$ degrees of freedom.

As explained in the main text, the differences moment conditions in (A.5) are a linear combination of the generic moment conditions in (A.4). The associated $q\times p$ transformation matrix of rank q that transforms (A.4) into (A.5) is denoted by $\mathbf{M}_1$. We write (A.4) as $\mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})]=\mathbf{0}$ and (A.5) as $\mathbf{M}_1\mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})]=\mathbf{0}$.

Let $\mathbf{d}\neq\mathbf{0}$ be a local alternative such that the within conditions in (A.3) do not hold, implying that (A.4) and (A.5) cannot hold either. This means that we must have $\mathbf{M}_1\mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})]=\mathbf{d}_1/\sqrt{n}\neq\mathbf{0}$. This is Case 1 in Table A.1, where the threefold 'no' indicates that none of the three population moment conditions hold under the local alternative. The non-centrality parameter of the J_1-test for this local alternative is denoted $\lambda_1(\mathbf{d}_1)$. The non-centrality parameter of the J_2-test is denoted by $\lambda_2(\mathbf{d})$. Because the J_2-test has the maximum number of degrees of freedom, its non-centrality parameter has the largest possible value $\lambda^*(\mathbf{d})$ for this local alternative (Newey, 1985, Prop. 6). Hence $\lambda_1(\mathbf{d}_1)\leq\lambda_2(\mathbf{d})=\lambda^*(\mathbf{d})$.

Now take a $(p-q)\times p$ matrix $\mathbf{M}_2$ of rank $p-q$ such that $\mathbf{L}=[\mathbf{M}_1'\ \mathbf{M}_2']'$ has rank p. The J-test for the moment conditions $\mathbf{L}\mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})]=\mathbf{0}$ has the maximum number of degrees of freedom and must therefore also have the maximum non-centrality parameter $\lambda^*(\mathbf{d})$. We conclude that $\lambda^*(\mathbf{d})$ cannot depend on $\mathbf{d}_2=\mathbf{M}_2\mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})]$; otherwise there would exist a J-test with the maximum number of degrees of freedom but a higher or lower non-centrality parameter than $\lambda^*(\mathbf{d})$. We must therefore have $\lambda^*(\mathbf{d})=\lambda^*(\mathbf{d}_1)$.

Let $\tilde{\mathbf{d}}$ be a local alternative with $\mathbf{M}_1\mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})]=\mathbf{d}_1/\sqrt{n}\neq\mathbf{0}$, but $\mathbf{M}_2\mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})]=\mathbf{0}$. Because $\lambda^*(\mathbf{d}_1)$ only depends on $\mathbf{d}_1$, we must have $\lambda_2(\tilde{\mathbf{d}})=\lambda_2(\mathbf{d})=\lambda^*(\mathbf{d}_1)$. Furthermore, because $\mathbf{M}_2\mathbb{E}[\mathbf{g}_i(\boldsymbol{\beta})]=\mathbf{0}$, we must also have $\lambda_2(\mathbf{d}_1)=\lambda_1(\mathbf{d}_1)$. We thus conclude that $\lambda^*(\mathbf{d}_1)=\lambda_1(\mathbf{d}_1)=\lambda_2(\mathbf{d}_1)$, showing that the J_1- and J_2-tests both have the same maximum value of the non-centrality parameter for local alternatives such that (A.3) does not hold.

Because the J_1-test has a lower number of degrees of freedom, its local power is higher than that of the J_2-test for such alternatives (Newey, 1985, p. 238). Hence, the J_1-test for the differences moment conditions in (A.5) is the more powerful J-test to detect local alternatives for which (A.3) does not hold. Only if the non-centrality parameter turns out 0 for such alternatives – which may happen as shown by Newey (1985, Prop. 1) and Hall (2005, Th. 5.4) – the local power of both tests is the same and equal to α.

For local alternatives such that (A.3) does hold, there are three possible cases; see Cases 2–3–4 in Table A.1.[1] For each case, the asymptotic distributions of the two J-tests are given in Table A.1. In case of a central chi-square distribution, the J-test's local power is α (i.e., nominal). If the asymptotic distribution is non-central chi-square, the power exceeds α (above nominal). From Table A.1 it becomes clear that each J-test's local power is nominal or above nominal whenever (A.3) holds, making both tests conservative. □

Table A.1: Asymptotic distribution of J-test statistics under different local alternatives

case	within	differences	generic	J_1	J_2
1	no	no	no	$\chi^2_{c_1,q-k}$	$\chi^2_{c_1,p-k}$
2	yes	yes	yes	χ^2_{q-k}	χ^2_{p-k}
3	yes	yes	no	$\chi^2_{c_3,q-k}$	χ^2_{p-k}
4	yes	no	no	$\chi^2_{c_4,q-k}$	$\chi^2_{c_4,p-k}$

Notes: This table compares two J-tests: for the generic conditions (J_1) and the differences conditions (J_2). The yes/no in each case refers to the moment conditions that apply under the local alternative. The within conditions are given in (A.3), the differences conditions in (A.5) and the generic conditions in (A.4). The central chi-square distribution with degrees of freedom df is denoted by χ^2_{df}. The non-central chi-square distribution with non-centrality parameter c and degrees of freedom df is denoted by $\chi^2_{c,df}$. The non-centrality parameter is numbered for each case to emphasize its dependence on the local alternative. The expression for the non-centrality parameter is given in (C.30).

B. Motivating examples: calculations

Properties of the AR(1) model

This appendix makes use of a few elementary properties of stationary AR(1) processes, which we summarize here for completeness. Assume that x_{it} and z_{it} are generated by stationary AR(1) processes, such that

$$x_{it} = \rho x_{i,t-1} + \theta_{it} \qquad [0 < \rho < 1] \tag{B.1}$$

$$z_{it} = \delta z_{i,t-1} + \eta_{it} \qquad [0 < \delta < 1]. \tag{B.2}$$

[1]Because (A.4) $\Longrightarrow$ (A.5) $\Longrightarrow$ (A.3) not all combinations of moment conditions are possible, which explains why Table A.1 considers four cases in total.

We assume that $\mathbb{E}(\theta_{it}) = \mathbb{E}(\eta_{it}) = 0$, $\mathbb{E}(\theta_{it}^2) = \sigma_\theta^2$ and $\mathbb{E}(\eta_{it}^2) = \sigma_\eta^2$ for all i and t. We also assume that $\mathbb{C}\mathrm{ov}\,(\theta_{mt}, \eta_{is}) = 0$ for $m \neq n$, $\mathbb{C}\mathrm{ov}\,(\theta_{is}, \eta_{it}) = 0$ for $s \neq t$, and $\mathbb{C}\mathrm{ov}\,(\theta_{it}, \eta_{it}) = \sigma_{\theta\eta}$. Lastly, we assume that $\mathbb{C}\mathrm{ov}\,(\theta_{mt}, \varepsilon_{is}) = \mathbb{C}\mathrm{ov}\,(\eta_{mt}, \varepsilon_{is}) = 0$ for all m, i, s, t.

For $k \geq 1$, we can write

$$x_{it} = \rho^k x_{i,t-k} + \sum_{\ell=0}^{k-1} \rho^\ell \theta_{i,t-\ell}, \qquad z_{it} = \delta^k z_{i,t-k} + \sum_{\ell=0}^{k-1} \delta^\ell \eta_{i,t-\ell}. \tag{B.3}$$

By letting $k \to \infty$, we find

$$x_{it} = \sum_{\ell=0}^{\infty} \rho^\ell \theta_{i,t-\ell}, \qquad z_{it} = \sum_{\ell=0}^{\infty} \delta^\ell \eta_{i,t-\ell}. \tag{B.4}$$

Using these alternative formulations for x_{it} and z_{it}, we find for $j \geq 0$,

$$\mathbb{V}\mathrm{ar}\,(x_{it}) = \sigma_\theta^2/(1-\rho^2) \equiv \sigma_x^2, \quad \mathbb{V}\mathrm{ar}\,(z_{it}) = \sigma_\eta^2/(1-\delta^2) \equiv \sigma_z^2, \tag{B.5}$$

$$\mathbb{C}\mathrm{ov}\,(x_{it}, x_{i,t-j}) = \rho^j \sigma_x^2, \quad \mathbb{C}\mathrm{ov}\,(z_{it}, z_{i,t-j}) = \delta^j \sigma_z^2. \tag{B.6}$$

We also have

$$\mathbb{C}\mathrm{ov}\,(x_{it}, z_{i,t-j}) = \delta^{-j} \sum_{\ell=j}^{\infty} (\delta\rho)^\ell \mathbb{C}\mathrm{ov}\,(\theta_{i,t-\ell}, \eta_{i,t-\ell}) = \delta^{-j} \sigma_{\theta\eta} \sum_{\ell=j}^{\infty} (\delta\rho)^\ell = \rho^j \sigma_{\theta\eta}/(1-\delta\rho). \tag{B.7}$$

Similarly, we find

$$\mathbb{C}\mathrm{ov}\,(x_{i,t-j}, z_{it}) = \delta^j \sigma_{\theta\eta}/(1-\delta\rho). \tag{B.8}$$

B.1. (Non-)Classical measurement error

We start with the errors-in-variables model and allow for non-classical measurement error, with classical measurement as a special case. We will derive the inconsistency in both cases.

Model. Consider the linear panel regression model with measurement error, given by

$$y_{it} = \alpha_i + \beta \xi_{it} + \varepsilon_{it} \tag{B.9}$$

$$x_{it} = \xi_{it} + v_{it}, \tag{B.10}$$

where $n = 1, \ldots, n$ and $t = 1, \ldots, T$. We assume that (ε_{it}) is i.i.d. with $\mathbb{E}(\varepsilon_{it}) = 0$ and $\mathbb{E}(\varepsilon_{it}^2) = \sigma_\varepsilon^2$ for all i and t. Regarding (ξ_{it}) and (v_{it}), we assume that they are generated by stationary AR(1) processes, such that

$$\xi_{it} = \rho\xi_{i,t-1} + \theta_{it} \qquad [0 < \rho < 1] \tag{B.11}$$

$$v_{it} = \delta v_{i,t-1} + \eta_{it} \qquad [0 < \delta < 1]. \tag{B.12}$$

We assume that $\mathbb{E}(\theta_{it}) = \mathbb{E}(\eta_{it}) = 0$, $\mathbb{E}(\theta_{it}^2) = \sigma_\theta^2$ and $\mathbb{E}(\eta_{it}^2) = \sigma_\eta^2$ for all i and t. Furthermore, we assume that $\mathbb{C}\text{ov}\,(\theta_{mt}, \eta_{is}) = 0$ for $m \neq n$, $\mathbb{C}\text{ov}\,(\theta_{is}, \eta_{it}) = 0$ for $s \neq t$, $\mathbb{C}\text{ov}\,(\theta_{it}, \eta_{it}) = \sigma_{\theta\eta}$ and $\mathbb{C}\text{ov}\,(\theta_{mt}, \varepsilon_{is}) = 0$ for all m, i, s, t. Lastly, we assume that $\mathbb{C}\text{ov}\,(\varepsilon_{mt}, \eta_{is}) = 0$ for $m \neq i$, $\mathbb{C}\text{ov}\,(\varepsilon_{is}, \eta_{it}) = 0$ for all s, t. If $\sigma_{\theta\eta} \neq 0$, we have a form of non-classical measurement error.

Inconsistency. We first show that the within estimator will usually be inconsistent. Let $\mathbf{x}_i = (x_{i1}, \ldots, x_{iT})'$ and $\mathbf{y}_i = (y_{i1}, \ldots, y_{iT})'$. With $\mathbf{A}$ the $T \times T$ centering matrix we obtain

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}_w = \operatorname*{plim}_{n\to\infty} \frac{\sum_i \mathbf{x}_i'\mathbf{A}\mathbf{y}_i}{\sum_i \mathbf{x}_i'\mathbf{A}\mathbf{x}_i} = \frac{\text{tr}[\mathbf{A}(\mathbf{\Sigma}_\xi + \mathbf{\Sigma}_{\xi v})]}{\text{tr}[\mathbf{A}(\mathbf{\Sigma}_\xi + \mathbf{\Sigma}_v)]}\beta, \tag{B.13}$$

where $\mathbf{\Sigma}_v$ contains the covariances $\mathbb{C}\text{ov}\,(v_{ns}, v_{nt})$ and $\mathbf{\Sigma}_{\xi v}$ the covariances $\mathbb{C}\text{ov}\,(\xi_{ns}, v_{nt})$. The probability limit will typically be unequal to β if at least $\mathbf{\Sigma}_v \neq \mathbf{0}$.

We now turn to the estimators $\hat{\beta}_j$ that are obtained after taking differences over time span j. It holds that

$$\begin{aligned}
\operatorname*{plim}_{n\to\infty} \hat{\beta}_j &= \frac{\mathbb{C}\text{ov}\,(y_{it} - y_{i,t-j}, x_{it} - x_{i,t-j})}{\mathbb{V}\text{ar}\,(x_{it} - x_{i,t-j})} \\
&= \frac{\mathbb{C}\text{ov}\,(\beta(x_{it} - x_{i,t-j}) - \beta(v_{it} - v_{i,t-j}) + \varepsilon_{it} - \varepsilon_{i,t-j}, x_{it} - x_{i,t-j})}{\mathbb{V}\text{ar}\,(x_{it} - x_{i,t-j})} \\
&= \beta + \frac{\mathbb{C}\text{ov}\,(\varepsilon_{it} - \varepsilon_{i,t-j}, x_{it} - x_{i,t-j}) - \beta\mathbb{C}\text{ov}\,(v_{it} - v_{i,t-j}, x_{it} - x_{i,t-j})}{\mathbb{V}\text{ar}\,(x_{it} - x_{i,t-j})}.
\end{aligned} \tag{B.14}$$

Under the given assumptions, the numerator in (B.14) reduces to

$$\begin{aligned}
&\mathbb{C}\text{ov}\,(\varepsilon_{it} - \varepsilon_{i,t-j}, v_{it} - v_{i,t-j}) - \beta[\mathbb{V}\text{ar}\,(v_{it} - v_{i,t-j}) + \mathbb{C}\text{ov}\,(v_{it} - v_{i,t-j}, \xi_{it} - \xi_{i,t-j})] = \\
&-2\beta[\sigma_\eta^2(1 - \delta^j)/(1 - \delta^2) + \sigma_{\theta\eta}(1 - (\delta^j + \rho^j)/2)/(1 - \delta\rho)].
\end{aligned} \tag{B.15}$$

Furthermore, the denominator can be written as

$$\begin{aligned}
&\mathbb{V}\text{ar}\,(\xi_{it} - \xi_{i,t-j} + v_{it} - v_{i,t-j}) = \\
&\mathbb{V}\text{ar}\,(\xi_{it} - \xi_{i,t-j}) + \mathbb{V}\text{ar}\,(v_{it} - v_{i,t-j}) + 2\mathbb{C}\text{ov}\,(\xi_{it} - \xi_{i,t-j}, v_{it} - v_{i,t-j}) = \\
&2[\sigma_\theta^2(1 - \rho^j)/(1 - \rho^2) + \sigma_\eta^2(1 - \delta^j)/(1 - \delta^2) + 2\sigma_{\theta\eta}(1 - (\delta^j + \rho^j)/2)/(1 - \delta\rho)].
\end{aligned} \tag{B.16}$$

The inconsistency thus boils down to

$$\begin{aligned}\operatorname*{plim}_{n\to\infty}\hat{\beta}_j - \beta &= \frac{-\beta[\sigma_\eta^2(1-\delta^j)/(1-\delta^2) + \sigma_{\theta\eta}(1-(\delta^j+\rho^j)/2)/(1-\delta\rho)]}{\sigma_\theta^2(1-\rho^j)/(1-\rho^2) + \sigma_\eta^2(1-\delta^j)/(1-\delta^2) + 2\sigma_{\theta\eta}(1-(\delta^j+\rho^j)/2)/(1-\delta\rho)} \\ &= \frac{-\beta[\sigma_\nu^2(1-\delta^j) + \sigma_{\xi\nu}(1-(\delta^j+\rho^j)/2)]}{\sigma_\xi^2(1-\rho^j) + \sigma_\nu^2(1-\delta^j) + 2\sigma_{\xi\nu}(1-(\delta^j+\rho^j)/2)}. \end{aligned} \tag{B.17}$$

Because

$$(1-\delta^j)(1-\rho^{j+1}) > (1-\delta^{j+1})(1-\rho^j) \tag{B.18}$$

if and only if $\delta < \rho$, it is readily seen that the inconsistency's magnitude decreases with j if and only if $\delta < \rho$. For $\delta > \rho$, the magnitude of the inconsistency is increasing and for $\delta = \rho$ the inconsistency does not depend on j. For both classical and non-classical measurement error, the inconsistency does not vanish for larger values of j.

B.2. Omitted variables

The second source of endogeneity that we consider is an omitted variable.

Model. Consider the linear panel regression model with two regressors, given by

$$y_{it} = \alpha_i + \beta x_{it} + \gamma z_{it} + \varepsilon_{it}, \tag{B.19}$$

where $i = 1, \ldots, n$ and $t = 1, \ldots, T$. We assume that (ε_{it}) is i.i.d. with $\mathbb{E}(\varepsilon_{it}) = 0$ and $\mathbb{E}(\varepsilon_{it}^2) = \sigma_\varepsilon^2$ for all i and t. Regarding the explanatory variables, we assume that x_{it} and z_{it} are generated by stationary AR(1) processes, such that

$$\begin{aligned} x_{it} &= \rho x_{i,t-1} + \theta_{it} \qquad [0 < \rho < 1] \quad &(B.20)\\ z_{it} &= \delta z_{i,t-1} + \eta_{it} \qquad [0 < \delta < 1]. \quad &(B.21)\end{aligned}$$

We assume that $\mathbb{E}(\theta_{it}) = \mathbb{E}(\eta_{it}) = 0$, $\mathbb{E}(\theta_{it}^2) = \sigma_\theta^2$ and $\mathbb{E}(\eta_{it}^2) = \sigma_\eta^2$ for all i and t. Furthermore, we assume that $\mathbb{C}\mathrm{ov}\,(\theta_{mt}, \eta_{is}) = 0$ for $m \neq n$, $\mathbb{C}\mathrm{ov}\,(\theta_{is}, \eta_{it}) = 0$ for $s \neq t$, and $\mathbb{C}\mathrm{ov}\,(\theta_{it}, \eta_{it}) = \sigma_{\theta\eta}$. Lastly, we assume that $\mathbb{C}\mathrm{ov}\,(\theta_{mt}, \varepsilon_{is}) = \mathbb{C}\mathrm{ov}\,(\eta_{mt}, \varepsilon_{is}) = 0$ for all m, n, s, t.

We estimate the omitted-variable regression

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it}, \tag{B.22}$$

and are interested in the probability limit of $\hat{\beta}_j$, the estimator of β based on the model after taking differences over time span j.

Inconsistency. We first show that the within estimator for β will usually be inconsistent. Using similar matrix notation as before, we obtain

$$\operatorname*{plim}_{n\to\infty}\hat{\beta}_w = \operatorname*{plim}_{n\to\infty}\frac{\sum_i \mathbf{x}_i'\mathbf{A}\mathbf{y}_i}{\sum_i \mathbf{x}_i'\mathbf{A}\mathbf{x}_i} = \beta + \frac{\operatorname{tr}(\mathbf{A}\mathbf{\Sigma}_{zx})\gamma}{\operatorname{tr}(\mathbf{A}\mathbf{\Sigma}_{x})}, \tag{B.23}$$

which will be unequal to β for $\gamma \neq 0$ and $\mathbf{\Sigma}_{zx} \neq \mathbf{0}$.

We now turn to the estimators $\hat{\beta}_j$ that are obtained after taking differences over time span j. It holds that

$$\operatorname*{plim}_{n\to\infty}\hat{\beta}_j = \frac{\mathbb{C}\mathrm{ov}\,(y_{it}-y_{i,t-j}, x_{it}-x_{i,t-j})}{\mathbb{V}\mathrm{ar}\,(x_{it}-x_{i,t-j})}. \tag{B.24}$$

Under the given assumptions, the numerator reduces to

$$\begin{aligned}
&\mathbb{C}\mathrm{ov}\,(y_{it}-y_{i,t-j}, x_{it}-x_{i,t-j}) = \beta\mathbb{V}\mathrm{ar}\,(x_{it}-x_{i,t-j}) + \gamma\mathbb{C}\mathrm{ov}\,(x_{it}-x_{i,t-j}, z_{it}-z_{i,t-j}) = \\
&\beta\mathbb{V}\mathrm{ar}\,(x_{it}-x_{i,t-j}) + \gamma[2\mathbb{C}\mathrm{ov}\,(x_{it}, z_{it}) - \mathbb{C}\mathrm{ov}\,(x_{i,t-j}, z_{it}) - \mathbb{C}\mathrm{ov}\,(x_{it}, z_{i,t-j})] = \\
&2[\beta\sigma_\theta^2(1-\rho^j)/(1-\rho^2) + \gamma\sigma_{\theta\eta}(1-(\delta^j+\rho^j)/2)/(1-\delta\rho)].
\end{aligned} \tag{B.25}$$

For the denominator, we find

$$\mathbb{V}\mathrm{ar}\,(x_{it}-x_{i,t-j}) = 2[\mathbb{V}\mathrm{ar}\,(x_{it}) - \mathbb{C}\mathrm{ov}\,(x_{it}, x_{i,t-j})] = 2\sigma_x^2(1-\rho^j) = 2\sigma_\theta^2(1-\rho^j)/(1-\rho^2). \tag{B.26}$$

The probability limit then becomes

$$\begin{aligned}
\operatorname*{plim}_{n\to\infty}\hat{\beta}_j &= \frac{\beta\sigma_\theta^2(1-\rho^j)/(1-\rho^2) + \gamma\sigma_{\theta\eta}(1-(\delta^j+\rho^j)/2)/(1-\delta\rho)}{\sigma_\theta^2(1-\rho^j)/(1-\rho^2)} \\
&= \beta + \frac{\gamma\sigma_{\theta\eta}[1-(\delta^j+\rho^j)/2]/(1-\delta\rho)}{\sigma_\theta^2(1-\rho^j)/(1-\rho^2)}.
\end{aligned} \tag{B.27}$$

The inconsistency thus boils down to

$$\operatorname*{plim}_{n\to\infty}\hat{\beta}_j - \beta = \frac{\gamma\sigma_{\theta\eta}[1-(\delta^j+\rho^j)/2]/(1-\delta\rho)}{\sigma_\theta^2(1-\rho^j)/(1-\rho^2)} = \frac{\gamma\sigma_{xz}[1-(\delta^j+\rho^j)/2]}{\sigma_x^2(1-\rho^j)}. \tag{B.28}$$

As a sanity check on the above expression, we notice that the inconsistency is zero for $\sigma_{\theta\eta} = 0$. The inconsistency should be zero in this particular case, because $\sigma_{\theta\eta} = 0$ implies that x_{it} and z_{it} are uncorrelated.

Because

$$(1-\rho^{j+1})(1-(\delta^{j}+\rho^{j})/2) > (1-\rho^{j})(1-(\delta^{j+1}+\rho^{j+1})/2) \tag{B.29}$$

if and only if $\delta < \rho$, is readily seen that $\text{plim}_{n\to\infty} |\hat{\beta}_j - \beta| > \text{plim}_{n\to\infty} |\hat{\beta}_{j+1} - \beta|$ if and only if $\delta < \rho$. The inconsistency's magnitude is increasing for $\delta > \rho$ and for $\delta = \rho$ the inconsistency does not depend on j. We note that the inconsistency does not vanish for larger values of j.

Extension to time-varying covariates. The general case of omitted variables encompasses the case of an ignored time-varying coefficient. Models with time-varying coefficients have been considered in production and cost analysis to deal with technical change that affects e.g. marginal costs and productivity growth (e.g., Koetter et al., 2012). To see the relation with omitted variables, consider the linear panel regression model with a single regressor and a coefficient that is a deterministic function of time, given by

$$y_{it} = \alpha_i + \beta(t)x_{it} + \varepsilon_{it}, \tag{B.30}$$

for $i = 1, \ldots, n$ and $t = 1, \ldots, T$. We consider the case that $\beta(t) = b_0 + b_1 t + b_2 t^2$ for scalars b_0, b_1, b_2. This functional form of the time-varying coefficient has also been used in the aforementioned production and cost literature.

We assume that (ε_{it}) is i.i.d. with $\mathbb{E}(\varepsilon_{it}) = 0$ and $\mathbb{E}(\varepsilon_{it}^2) = \sigma_\varepsilon^2$ for all i and t. Regarding the explanatory variable, we assume that x_{it} is generated by a stationary AR(1) process, such that

$$x_{it} \quad = \quad \rho x_{i,t-1} + \theta_{it} \qquad [0 < \rho < 1]. \tag{B.31}$$

We assume that $\mathbb{E}(\theta_{it}) = 0$ and $\mathbb{E}(\theta_{it}^2) = \sigma_\theta^2$ for all i and t. We also assume that $\mathbb{C}\text{ov}\,(\theta_{it}, \varepsilon_{js}) = 0$ for all i, j, s, t.

We create endogeneity by ignoring the possibility of a time-varying coefficient and estimate the misspecified regression with a time-constant β; i.e. we estimate

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it}. \tag{B.32}$$

We now see that the linear panel regression model in (B.32) contains two omitted variables, namely tx_{it} and $t^2 x_{it}$. Because this case of omitted variables is analytically hard to deal with, our calculations are less detailed as before.

Let $\tau_i = (x_{i1}, 2x_{i2}, \ldots, Tx_{iT})'$ and $\tilde{\tau}_i = (x_{i1}, 4x_{i2}, \ldots, T^2 x_{iT})'$. Using similar matrix notation as before, we obtain for the within estimator

$$\operatorname*{plim}_{n\to\infty}\hat{\beta}_w = \operatorname*{plim}_{n\to\infty}\frac{\sum_i \mathbf{x}_i'\mathbf{A}\mathbf{y}_i}{\sum_i \mathbf{x}_i'\mathbf{A}\mathbf{x}_i} = b_0 + \frac{\operatorname{tr}(\mathbf{A}\mathbf{\Sigma}_{x\tau})b_1 + \operatorname{tr}(\mathbf{A}\mathbf{\Sigma}_{x\tilde{\tau}})b_2}{\operatorname{tr}(\mathbf{A}\mathbf{\Sigma}_x)}, \tag{B.33}$$

which will typically be inconsistent for b_0 for $b_1 \neq 0$ or $b_2 \neq 0$. Similarly, we find

$$\operatorname*{plim}_{n\to\infty}\hat{\beta}_j = \operatorname*{plim}_{n\to\infty}\frac{\sum_i \mathbf{x}_i'\mathbf{\Delta}_j\mathbf{y}_i}{\sum_i \mathbf{x}_i'\mathbf{\Delta}_j\mathbf{x}_i} = b_0 + \frac{\operatorname{tr}(\mathbf{\Delta}_j\mathbf{\Sigma}_{x\tau})b_1 + \operatorname{tr}(\mathbf{\Delta}_j\mathbf{\Sigma}_{x\tilde{\tau}})b_2}{\operatorname{tr}(\mathbf{\Delta}_j\mathbf{\Sigma}_x)}. \tag{B.34}$$

Given (B.34), the inconsistency of $\operatorname{plim}_{n\to\infty}\hat{\beta}_j$ relative to b_0 is genereally expected to depend on j.

B.3. Simultaneity

The third source of endogeneity that we consider is simultaneity.

Model. We consider the simultaneous equations model given by the structural equations

$$y_{it} = \beta_i + \beta x_{it} + \varepsilon_{it} \tag{B.35}$$

$$x_{it} = \alpha_i + \alpha y_{it} + u_{it}. \tag{B.36}$$

We assume that (ε_{it}) is i.i.d. with $\mathbb{E}(\varepsilon_{it}) = 0$ and $\mathbb{V}\mathrm{ar}\,(\varepsilon_{it}) = \sigma_\varepsilon^2$, independent of (u_{it}). Here (u_{it}) is a stationary AR(1) process defined by

$$u_{it} = \rho u_{i,t-1} + \theta_{it} \qquad [0 < \rho < 1], \tag{B.37}$$

with $\mathbb{E}(\theta_{it}) = 0$, $\mathbb{E}(\theta_{it}^2) = \sigma_\theta^2$ and $\mathbb{C}\mathrm{ov}\,(\theta_{mt}, \varepsilon_{is}) = \mathbf{0}$ for all m, i, t, s.

Solving the two equations yields the reduced forms

$$y_{it} = \frac{\beta_i + \beta\alpha_i}{1-\alpha\beta} + \frac{\beta u_{it} + \varepsilon_{it}}{1-\alpha\beta} \tag{B.38}$$

$$x_{it} = \frac{\alpha_i + \alpha\beta_i}{1-\alpha\beta} + \frac{u_{it} + \alpha\varepsilon_{it}}{1-\alpha\beta}. \tag{B.39}$$

We estimate (B.35) in j-th differences, thereby ignoring (B.36). We are interested in the probability limit of $\hat{\beta}_j$, the estimator of β based on the model in j-th differences. We want to know how the inconsistency depends on j.

Inconsistency. We first show that the within estimator for β will usually be inconsistent. Using similar matrix notation as before, we obtain

$$\operatorname*{plim}_{n\to\infty}\hat{\beta}_w = \operatorname*{plim}_{n\to\infty}\frac{\sum_i \mathbf{x}_i'\mathbf{A}\mathbf{y}_i}{\sum_i \mathbf{x}_i'\mathbf{A}\mathbf{x}_i} = \frac{\operatorname{tr}(\mathbf{A}\mathbf{\Sigma}_u)\beta + \operatorname{tr}(\mathbf{A}\mathbf{\Sigma}_\varepsilon)\alpha}{\operatorname{tr}(\mathbf{A}\mathbf{\Sigma}_u) + \operatorname{tr}(\mathbf{A}\mathbf{\Sigma}_\varepsilon)\alpha^2}, \tag{B.40}$$

which will be unequal to β if $\alpha \neq 0$ and $\alpha \neq 1/\beta$.

We now turn to the estimators $\hat{\beta}_j$ that are obtained after taking differences over time span j. The probability limit of the resulting estimator for β equals

$$\begin{aligned}
\operatorname*{plim}_{n\to\infty} \hat{\beta}_j &= \frac{\mathbb{C}\mathrm{ov}\,(y_{it} - y_{i,t-j}, x_{it} - x_{i,t-j})}{\mathbb{V}\mathrm{ar}\,(x_{it} - x_{i,t-j})} \\
&= \frac{[\alpha \mathbb{V}\mathrm{ar}\,(\varepsilon_{it} - \varepsilon_{i,t-j}) + \beta \mathbb{V}\mathrm{ar}\,(u_{it} - u_{i,t-j}]/(1-\alpha\beta)^2}{\mathbb{V}\mathrm{ar}\,(x_{it} - x_{i,t-j})} \\
&= \frac{\alpha \mathbb{V}\mathrm{ar}\,(\varepsilon_{it} - \varepsilon_{i,t-j}) + \beta \mathbb{V}\mathrm{ar}\,(u_{it} - u_{i,t-j})}{\alpha^2 \mathbb{V}\mathrm{ar}\,(\varepsilon_{it} - \varepsilon_{i,t-j}) + \mathbb{V}\mathrm{ar}\,(u_{it} - u_{i,t-j})}.
\end{aligned} \tag{B.41}$$

Under the given assumption, this reduces to

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}_j = \frac{2[\alpha\sigma_\varepsilon^2 + \beta\sigma_\theta^2(1-\rho^j)/(1-\rho^2)]}{2[\alpha^2\sigma_\varepsilon^2 + \sigma_\theta^2(1-\rho^j)/(1-\rho^2)]} = \frac{\alpha\sigma_\varepsilon^2 + \beta\sigma_\theta^2(1-\rho^j)/(1-\rho^2)}{\alpha^2\sigma_\varepsilon^2 + \sigma_\theta^2(1-\rho^j)/(1-\rho^2)} = \frac{\alpha\sigma_\varepsilon^2 + \beta\sigma_u^2(1-\rho^j)}{\alpha^2\sigma_\varepsilon^2 + \sigma_u^2(1-\rho^j)}.$$

This gives the inconsistency

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}_j - \beta = \frac{\alpha\sigma_\varepsilon^2(1-\alpha\beta)}{\alpha^2\sigma_\varepsilon^2 + \sigma_u^2(1-\rho^j)}. \tag{B.42}$$

The inconsistency is positive for $\alpha(1-\alpha\beta) > 0$ and negative for $\alpha(1-\alpha\beta) < 0$. Its magnitude decreases with j for $0 < \rho < 1$. We note that the inconsistency does not vanish for larger values of j.

C. Local power: calculations

The notation in this section deviates slightly from the main text.

C.1. Wald test

We assume that we estimate the linear panel regression model

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it}, \tag{C.1}$$

for $i = 1, \ldots, n$ and $t = 1, \ldots, T$. Here y_{it} is the dependent variable, α_i is an individual effect, x_{it} is an observed covariate, β an unknown coefficient and ε_{it} an error term. We regress y_{it} on x_{it} after taking differences over time span j:

$$y_{it} - y_{i,t-j} = \beta(x_{it} - x_{i,t-j}) + \varepsilon_{it} - \varepsilon_{i,t-j} \qquad [t = j+1, \ldots, T]. \tag{C.2}$$

The corresponding OLS estimator of β is denoted by $\hat{\beta}_j$ for $j = 1, \ldots, T-1$. We focus on fixed-T and large-n asymptotics and define $\beta_j \equiv \operatorname{plim}_{n\to\infty} \hat{\beta}_j$.

Our Wald test is based on $(\hat{\beta}_1, \ldots, \hat{\beta}_{T-1})'$. We write $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{T-1})'$ and $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \ldots, \hat{\beta}_{T-1})'$. Let

$$\begin{aligned}
\mathbf{y}_{i1} &= (y_{i2} - y_{i1}, y_{i3} - y_{i2}, \ldots\ldots, y_{iT} - y_{i,T-1})' \\
\mathbf{y}_{i2} &= (y_{i3} - y_{i1}, y_{i4} - y_{i2}, \ldots, y_{iT} - y_{i,T-2})' \\
&\vdots \\
\mathbf{y}_{i,T-1} &= y_{iT} - y_{i1},
\end{aligned}$$

and let $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{i,T-1}$ be defined analogously. Next, let

$$\mathbf{y}_i = \begin{pmatrix} \mathbf{y}_{i1} \\ \mathbf{y}_{i2} \\ \vdots \\ \mathbf{y}_{i,T-1} \end{pmatrix}, \qquad \mathbf{X}_i = \begin{pmatrix} \mathbf{x}_{i1} & 0 & \ldots & 0 \\ 0 & \mathbf{x}_{i2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathbf{x}_{i,T-1} \end{pmatrix}.$$

Then

$$\hat{\boldsymbol{\beta}} = \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i' \mathbf{X}_i \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i' \mathbf{y}_i. \tag{C.3}$$

A cluster-robust consistent estimator of the asymptotic covariance matrix of $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ is given by

$$\hat{\boldsymbol{\Sigma}} = \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i' \mathbf{X}_i \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i' \mathbf{u}_i \mathbf{u}_i' \mathbf{X}_i \right) \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i' \mathbf{X}_i \right)^{-1}, \tag{C.4}$$

where $\mathbf{u}_i = \mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\beta}}$.

To test the null hypothesis $H_0 : \beta_j = \beta_{j+1}$ for all $j = 1, \ldots, T-2$ against the alternative hypothesis $H_1 : \beta_j \neq \beta_{j+1}$ for at least one $j = 1, \ldots, T-2$, we write the Wald test statistic as

$$q_{\mathrm{w}} = n \hat{\boldsymbol{\beta}}' \mathbf{B}_1' (\mathbf{B}_1 \hat{\boldsymbol{\Sigma}} \mathbf{B}_1')^{-1} \mathbf{B}_1 \hat{\boldsymbol{\beta}}, \tag{C.5}$$

where $\mathbf{B}_1$ is the $(T-2) \times (T-1)$ matrix

$$\mathbf{B}_1 = \begin{pmatrix} -1 & 1 & 0 & \ldots & 0 \\ 0 & -1 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & -1 & 1 \end{pmatrix}. \tag{C.6}$$

Under the null hypothesis, q_{w} is asymptotically χ^2_{T-2} distributed. The Wald test is consistent against any fixed alternative $H_1 : \mathbf{B}_1\boldsymbol{\beta} \neq 0$. Under a local alternative $H_1 : \mathbf{B}_1\boldsymbol{\beta} = \mathbf{d}/\sqrt{n}$, the Wald test statistic is non-central chi-square distributed with $T - 2$ degrees of freedom and non-centrality parameter $c = \mathbf{d}'(\mathbf{B}_1\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1}\mathbf{B}_1')^{-1}\mathbf{d}$, with $\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1} = \operatorname{plim}_{\substack{n\to\infty \\ H_1}} \hat{\mathbf{\Sigma}}$; see Cameron and Trivedi (2005, Ch. 7).

C.2. Local power of Wald test: $T = 3$

Preliminaries for the case $T = 3$. We start with the case of $T = 3$ and will generalize to higher values of T later. We consider the errors-in-variables model with classical measurement error and assume that η_{it} is normally distributed with mean 0 and variance $\sigma_\eta^2 = \tilde{\sigma}_\eta^2/\sqrt{n}$. Under this local alternative of the Pitman drift form, it holds that

$$\operatorname*{plim}_{\substack{n\to\infty \\ H_1}} \sqrt{n}(\hat{\beta}_2 - \hat{\beta}_1) = d_{\mathrm{CME}}, \qquad \operatorname*{plim}_{\substack{n\to\infty \\ H_1}} \hat{\beta}_1 = \operatorname*{plim}_{\substack{n\to\infty \\ H_1}} \hat{\beta}_2 = \beta. \tag{C.7}$$

Using (B.17) with $\sigma_{\theta\eta} = 0$, we find

$$d_{\mathrm{CME}} = \operatorname*{plim}_{\substack{n\to\infty \\ H_1}} \sqrt{n}(\hat{\beta}_2 - \hat{\beta}_1) = \frac{\beta\tilde{\sigma}_\eta^2(\rho - \delta)}{\sigma_\theta^2(1 + \delta)}. \tag{C.8}$$

The non-centrality parameter of the $\chi^2_{c,1}$ distribution under the local alternative is given by

$$c_{\mathrm{CME}} = d^2_{\mathrm{CME}}(\mathbf{B}_1\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1}\mathbf{B}_1')^{-1}, \tag{C.9}$$

with $\mathbf{B}_1 = [\ -1 \quad 1\ ]$. The term $\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1}$ enters the non-centrality parameter as $\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1} = \operatorname{plim}_{\substack{n\to\infty \\ H_1}} \hat{\mathbf{\Sigma}}$. The matrices $\mathbf{G}$ and $\mathbf{H}$ will be specified below.

Asymptotic covariance matrix. The asymptotic covariance matrix of $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ is estimated as

$$\hat{\mathbf{\Sigma}} = \left(\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i'\mathbf{X}_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i'\mathbf{u}_i\mathbf{u}_i'\mathbf{X}_i\right)\left(\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i'\mathbf{X}_i\right)^{-1} \tag{C.10}$$

where $\mathbf{u}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$.

To calculate the non-centrality parameter of the $\chi^2_{c,1}$ distribution, we need $\operatorname{plim}_{\substack{n\to\infty \\ H_1}} \hat{\mathbf{\Sigma}}$. This asymptotic covariance matrix is given by $\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1}$. Here $\mathbf{G}$ is the matrix

$$\mathbf{G} = \operatorname*{plim}_{\substack{n\to\infty \\ H_1}} \left(\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i'\mathbf{X}_i\right) = \begin{pmatrix} g_{11} & 0 \\ 0 & g_{22} \end{pmatrix},$$

with

$$g_{11} = \mathbb{E}\Big((\xi_{i2} - \xi_{i1})^2 + (\xi_{i3} - \xi_{i2})^2\Big) = \frac{4\sigma_\theta^2}{1+\rho} \tag{C.11}$$

and

$$g_{22} = \mathbb{E}\Big((\xi_{i3} - \xi_{i1})^2\Big) = 2\sigma_\theta^2. \tag{C.12}$$

The matrix $\mathbf{H}$ is given by

$$\mathbf{H} = \operatorname*{plim}_{\substack{n\to\infty \\ H_1}} \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i'\mathbf{u}_i\mathbf{u}_i'\mathbf{X}_i\right) = \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix},$$

where $\mathbf{u}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$. Omitting the terms from $\mathbf{H}$ that have a zero probability limit under the local alternative, we find

$$\begin{aligned}
h_{11} &= \mathbb{E}\Big((\xi_{i2} - \xi_{i1})^2(\varepsilon_{i2} - \varepsilon_{i1})^2 + (\xi_{i3} - \xi_{i2})^2(\varepsilon_{i3} - \varepsilon_{i2})^2 + 2(\xi_{i2} - \xi_{i1})(\xi_{i3} - \xi_{i2})(\varepsilon_{i2} - \varepsilon_{i1})(\varepsilon_{i3} - \varepsilon_{i2})\Big) \\
&= 2\left(\frac{4\sigma_\varepsilon^2\sigma_\theta^2}{1+\rho} + \frac{(1-\rho)\sigma_\theta^2\sigma_\varepsilon^2}{1+\rho}\right). && \text{(C.13)} \\
h_{12} &= \mathbb{E}\Big((\xi_{i2} - \xi_{i1})(\xi_{i3} - \xi_{i1})(\varepsilon_{i2} - \varepsilon_{i1})(\varepsilon_{i3} - \varepsilon_{i1}) + (\xi_{i3} - \xi_{i2})(\xi_{i3} - \xi_{i1})(\varepsilon_{i3} - \varepsilon_{i2})(\varepsilon_{i3} - \varepsilon_{i1})\Big) \\
&= 2\sigma_\varepsilon^2\sigma_\theta^2. && \text{(C.14)} \\
h_{22} &= \mathbb{E}\Big((\xi_{i3} - \xi_{i1})^2(\varepsilon_{i3} - \varepsilon_{i1})^2\Big) \\
&= 4\sigma_\theta^2\sigma_\varepsilon^2. && \text{(C.15)}
\end{aligned}$$

Key steps. Throughout, our derivations of $\mathbf{G}$ and $\mathbf{H}$ make use of three key results:

(1) η_{it}, θ_{it} and ε_{it} are mutually independent and each have zero means.

(2) Under H_1, $\sigma_\eta^2 = \tilde{\sigma}_\eta^2/\sqrt{n}$ and

$$\operatorname*{plim}_{\substack{n\to\infty \\ H_1}} \hat{\beta}_1 = \operatorname*{plim}_{\substack{n\to\infty \\ H_1}} \hat{\beta}_2 = \beta. \tag{C.16}$$

(3) Under normality of the mutually independent random variables x and y (Wansbeek and Meijer, 2000, p. 366), it holds that $\mathbb{E}(x^2y^2) = \mathbb{E}(x^2)\mathbb{E}(y^2) + 2[\mathbb{E}(xy)]^2$. We use this result to calculate

the expectations

$$\mathbb{E}\Big((v_{it}-v_{i,t-1})^2(v_{it}-v_{i,t-2})^2\Big)=\mathbb{E}\Big((v_{i,t-1}-v_{i,t-2})^2(v_{it}-v_{i,t-2})^2\Big), \tag{C.17}$$

$$\mathbb{E}\Big((v_{it}-v_{i,t-1})^2(v_{i,t-1}-v_{i,t-2})^2\Big), \tag{C.18}$$

$$\mathbb{E}\Big((v_{it}-v_{i,t-1})^4\Big),\quad \mathbb{E}\Big((v_{it}-v_{i,t-2})^4\Big). \tag{C.19}$$

Together, these results ensure that various terms vanish in the probability limit under the local alternative. For example, consider the term

$$\frac{1}{n}\sum_{i=1}^{n}\Big\{(v_{it}-v_{i,t-1})^2(v_{it}-v_{i,t-2})^2\Big\}, \tag{C.20}$$

that appears in the calculation of $\mathbf{H}$. After application of the third result, we see that this term vanishes in the probability limit under the local alternative, since

$$\frac{1}{n}\cdot n\cdot\frac{\tilde{\sigma}_\eta^4}{n}\frac{2(3+\delta)}{(1+\delta)}\rightarrow 0. \tag{C.21}$$

Non-centrality parameter. The non-centrality parameter of the $\chi^2_{c,1}$ distribution under the local alternative is given by

$$c_{\text{CME}}=d^2_{\text{CME}}(\mathbf{B}_1\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1}\mathbf{B}_1')^{-1}, \tag{C.22}$$

with $\mathbf{B}_1=[\ -1\quad 1\ ]$. This yields

$$c_{\text{CME}}=\frac{8\beta^2\tilde{\sigma}_\eta^4(\delta-\rho)^2}{\sigma_\theta^2\sigma_\varepsilon^2(1+\delta)^2(9-\rho^2)}. \tag{C.23}$$

Omitted variables and simultaneity. Following the same approach, we derive the non-centrality parameter in case of omitted variables and simultaneity. In both cases, we assume that η_{it} is normally distributed. For the model with an omitted variable (OV), the local alternative with Pitman drift is $\gamma=\tilde{\gamma}/\sqrt{n}$ and for simultaneity (S) it is $\alpha=\tilde{\alpha}/\sqrt{n}$. This results in

$$d_{\text{OV}}=\frac{\tilde{\gamma}\sigma_{\theta\eta}(1-\delta)(\delta-\rho)}{2\sigma_\theta^2(1-\delta\rho)}, \tag{C.24}$$

$$d_{\text{S}}=-\frac{\tilde{\alpha}\rho\sigma_\varepsilon^2}{\sigma_\theta^2}, \tag{C.25}$$

and

$$
\begin{aligned}
c_{\text{OV}} &= \frac{2\tilde{\gamma}^2\sigma^2_{\theta\eta}\sigma^2_{\theta}(1-\delta)^2(\delta-\rho)^2}{\sigma^2_{\varepsilon}\sigma^4_{\eta}(1-\delta\rho)^2(9-\rho^2)}, && \text{(C.26)}\\
c_{\text{S}} &= \frac{8\tilde{\alpha}^2\rho^2\sigma^2_{\varepsilon}}{\sigma^2_{\theta}(9-\rho^2)}. && \text{(C.27)}
\end{aligned}
$$

Because the case of non-classical measurement error involves elaborate expressions, we focus on the case of $\sigma_{\varepsilon\eta} = 0$ for simplicity. As before, we assume that η_{it} is normally distributed. The local alternative is $\sigma^2_{\eta} = \tilde{\sigma}^2_{\eta}/\sqrt{n}; \sigma_{\theta\eta} = \tilde{\sigma}_{\theta\eta}/\sqrt{n}$. We then find

$$
d_{\text{NCME}} = \frac{\beta(\rho-\delta)[2\tilde{\sigma}^2_{\eta}(1-\delta\rho)+\tilde{\sigma}_{\theta\eta}(1-\delta^2)]}{2\sigma^2_{\theta}(1+\delta)(1-\delta\rho)}. \tag{C.28}
$$

This yields

$$
c_{\text{NCME}} = \frac{2\beta^2(\delta-\rho)^2[\tilde{\sigma}_{\theta\eta}(1-\delta^2)+2\tilde{\sigma}^2_{\eta}(1-\delta\rho)]^2}{\sigma^2_{\varepsilon}\sigma^2_{\theta}(1+\delta^2)(1-\delta\rho)^2(9-\rho^2)}. \tag{C.29}
$$

For a general discussion of the local power of the Wald test, we refer to Cameron and Trivedi (2005, Ch. 7). We also note that normality is a sufficient but not a necessary condition. Under weaker conditions, the same results can be obtained.

C.3. Local power of Wald test: $T > 3$

Classical measurement error. We consider $T > 3$ for the case of classical measurement error. Let $\sigma^2_{\eta} = \tilde{\sigma}^2_{\eta}/\sqrt{n}$ under the local alternative, implying that also $\sigma^2_{\nu} = \tilde{\sigma}^2_{\nu}/\sqrt{n}$. We maintain the assumption that η_{it} is normally distributed. The non-centrality parameter of the $\chi^2_{c,T-2}$ distribution under the local alternative is

$$
c_{\text{CME}} = \mathbf{d}'_{\text{CME}}(\mathbf{B}_1\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1}\mathbf{B}'_1)^{-1}\mathbf{d}_{\text{CME}}, \tag{C.30}
$$

with $\mathbf{B}_1$ as given in (C.6). Furthermore, $\mathbf{d}_{\text{CME}}$ is a vector of length $T-2$, with the k-th element equal to

$$
-\beta\frac{\tilde{\sigma}^2_{\nu}}{\sigma^2_{\xi}}\left(\frac{1-\delta^{k+1}}{1-\rho^{k+1}} - \frac{1-\delta^k}{1-\rho^k}\right) \tag{C.31}
$$

$\mathbf{G}$ is the $(T-1)\times(T-1)$ diagonal matrix with the k-th diagonal element equal to

$$
g_k = -2\sigma^2_{\xi}(T-k)(1-\rho^k). \tag{C.32}
$$

Lastly, $\mathbf{H}$ is the $(T-1)\times(T-1)$ symmetric matrix with diagonal elements

$$h_{kk} = \sigma_\xi^2\sigma_\varepsilon^2\Big[4(T-k)(1-\rho^k) + 2(T-k-1)(\rho^{2k} - 2\rho^k + 1)\Big], \tag{C.33}$$

and off-diagonal elements (for $\ell > k$)

$$\begin{aligned} h_{k\ell} &= \sum_{s=k+1}^{T}\sum_{t=\ell+1}^{T}\Big\{(r_{|s-t|} - r_{|s-t+\ell|} - r_{|s-t-k|} + r_{|s-t+\ell-k|})[r^\varepsilon_{|s-t|} - r^\varepsilon_{|s-t+\ell|} - r^\varepsilon_{|s-t-k|} + r^\varepsilon_{|s-t+\ell-k|}]\Big\} && \text{(C.34)}\\ &= 2\sigma_\varepsilon^2(T-\ell)[r_0 - r_\ell - r_k + r_{\ell-k}] && \text{(C.35)}\\ &- \sigma_\varepsilon^2(r_k - r_{k+\ell} - r_0 + r_\ell)[(T-2\ell+1)\mathbb{1}_{\{\ell\le T/2\}} + (T-k-\ell)\mathbb{1}_{\{\ell+k+1\le T\}}]. && \text{(C.36)} \end{aligned}$$

Here $r_{|k|} = \rho^k\sigma_\xi^2 = \rho^k\sigma_\theta^2/(1-\rho^2)$ and $r^\varepsilon_{|k|} = \sigma_\varepsilon^2\mathbb{1}_{\{k=0\}}$.

Notice that we can write

$$\mathbf{d}_{\text{CME}} = \frac{\beta\tilde\sigma_\eta^2}{\sigma_\theta^2}\boldsymbol{\delta}^*_{\text{CME}}(\delta,\rho,T), \qquad \mathbf{G} = \sigma_\theta^2\mathbf{G}^*(\rho,T), \qquad \mathbf{H} = \sigma_\theta^2\sigma_\varepsilon^2\mathbf{H}^*(\rho,T), \tag{C.37}$$

where $\boldsymbol{\delta}^*_{\text{CME}}(\delta,\rho,T)$, $\mathbf{G}^*(\rho,T)$ and $\mathbf{H}^*(\rho,T)$ are (matrix) functions of δ and/or ρ and T. Consequently, we can also write

$$c_{\text{CME}} = \mathbf{d}'_{\text{CME}}(\mathbf{B}_1\mathbf{G}^{-1}\mathbf{H}\mathbf{G}^{-1}\mathbf{B}'_1)^{-1}\mathbf{d}_{\text{CME}} = \frac{\beta^2\tilde\sigma_\eta^4}{\sigma_\theta^2\sigma_\varepsilon^2}c^*_{\text{CME}}(\delta,\rho,T), \tag{C.38}$$

where $c^*_{\text{CME}}(\delta,\rho,T)$ is a function of δ,ρ and T only. Hence, $\mathbf{d}_{\text{CME}}$ and c^*_{CME} are 'separable' in a component that depends on δ, ρ and T only and a component that depends only on the remaining model parameters.

Omitted variables, simultaneity and non-classical measurement error. Proceeding in a similar fashion (maintaining the normality assumptions made in Appendix C.2), we find similar separability in case of omitted variables, simultaneity and non-classical measurement error (with $\sigma_{\varepsilon\eta} = 0$ but $\sigma_{\theta\eta} \neq 0$):

$$\begin{aligned} c_{\text{OV}} &= \frac{\tilde\gamma^2\sigma_{\theta\eta}^2\sigma_\theta^2}{\sigma_\eta^4\sigma_\varepsilon^2}c^*_{\text{OV}}(\delta,\rho,T), && \text{(C.39)}\\ c_{\text{S}} &= \frac{\tilde\alpha^2\sigma_\varepsilon^2}{\sigma_\theta^2}c^*_{\text{S}}(\rho,T), && \text{(C.40)}\\ c_{\text{NCME}} &= \frac{\beta^2}{\sigma_\theta^2\sigma_\varepsilon^2}c^*_{\text{NCME}}(\tilde\sigma_\eta,\tilde\sigma_{\theta\eta},\delta,\rho,T). && \text{(C.41)} \end{aligned}$$

We note that the local power decreases with the number of degrees of freedom (which is a function of T) and increases with c (Newey, 1985). Hence, even if we were able to show that c increases with T, the effect of T on the local power would remain ambiguous.

D. Simulation study

We run a simulation study to explore the finite-sample properties of the consistent Wald test in terms of power and size under a fixed alternative. We simulate the models for classical measurement error, omitted variables and simultaneity as considered in Table 1 in the main text, for both $T = 5$ and $T = 10$. For classical measurement error and omitted variables, we simulate two scenarios in terms of the persistence parameters: $\delta = 0.3$ and $\rho = 0.9$ and $\delta = 0.3$ and $\rho = 0.6$. For simultaneity, we consider the cases $\rho = 0.6$ and $\rho = 0.9$.

Tables D.1 and D.2 report, respectively, the empirical power and size of the Wald test in (9). The notes of Table D.1 report the full set of parameters for each model. This part of the table also reports the probability limits of the underlying model's R^2 and the reliability and noise-to-signal ratios for the models with classical measurement error. The simulation results show that the Wald test's power can turn out relatively low for smaller values of n. This is especially the case if the distance between ρ and δ is relatively small. Table D.2 also points out that the size of the Wald test may be conservative for smaller values of n and T. In the other cases, the rejection rates are close to nominal.

The right-hand-side panels of Tables D.1 and D.2 report the empirical power and size of the J-test based on the generic moment conditions in (A.4). We have already proven that this test's power is lower than the above Wald test's power for local alternatives under which the within estimator's moment conditions do not hold. The simulation allows us to consider three fixed alternatives under which the within estimator is inconsistent. Both tests are consistent against these fixed alternatives. We therefore compare the two tests' finite-sample performance in terms of empirical power and size.

We observe that the Wald test always outperforms the generic J-test for $n = 100$. For $n = 500$ and $n = 1,000$ the differences are smaller but the Wald test still outperforms the J-test in case there is a power difference between the two tests, with exception of one case ($T = 10$, $\rho = 0.6$, OV). In terms of size, the generic J-test tends to produce empirical rejection rates below the nominal level, especially for smaller values of n and T. This confirms the conclusion drawn in the main text about the conservativeness of the generic J-test.

Table D.1: Simulation results: empirical power

	$n = 100$	$n = 500$	$n = 1,000$	$n = 100$	$n = 500$	$n = 1,000$
	conditions (A.5) ('differences')			*conditions (A.4) ('generic')*		
$\rho = 0.9$		$T = 5$			$T = 5$	
ME	1.00	1.00	1.00	0.00	1.00	1.00
OV	1.00	1.00	1.00	0.00	1.00	1.00
S	1.00	1.00	1.00	0.00	1.00	1.00
$\rho = 0.6$						
ME	0.61	1.00	1.00	0.00	0.68	0.99
OV	0.73	1.00	1.00	0.00	1.00	1.00
S	0.96	1.00	1.00	0.00	1.00	1.00
$\rho = 0.9$		$T = 10$			$T = 10$	
ME	0.60	1.00	1.00	0.13	0.97	1.00
OV	0.87	1.00	1.00	0.82	1.00	1.00
S	0.77	1.00	1.00	0.21	1.00	1.00
$\rho = 0.6$						
ME	0.22	0.70	0.96	0.04	0.36	0.72
OV	0.29	0.85	0.99	0.19	0.99	1.00
S	0.53	1.00	1.00	0.11	0.94	1.00

Notes: We consider the Wald form of the J-test for the moment conditions in (A.5), as well as the J-test for the generic moment conditions for strict exogeneity in (A.4). In all cases, the simulation results are based on 1,000 simulation runs. The empirical power is obtained as the fraction of the number of simulation runs in which the test rejects the null hypothesis. The simulated models correspond to three of the illustrative cases listed in Table 1 of the main text. Parameters for classical measurement error ('ME'): $\beta = 1$, $\sigma_\varepsilon^2 = 1$, $\sigma_\theta^2 = 1.44$, $\sigma_\eta^2 = 0.64$, $\delta = 0.3$. This yields reliabilities of 0.92 ($\rho = 0.9$) and 0.76 ($\rho = 0.6$), noise-to-signal ratios of 0.09 ($\rho = 0.9$) and 0.312 ($\rho = 0.6$) and probability limits for the model's R^2 of 0.88 ($\rho = 0.9$) and 0.69 ($\rho = 0.6$). Parameters for omitted variables ('OV'): $\beta = 1$, $\gamma = 1$, $\sigma_\varepsilon^2 = 0.25$, $\sigma_\theta^2 = 0.36$, $\sigma_\eta^2 = 0.36$, $\rho_{\theta\eta} = -0.6$, $\delta = 0.3$. This yields probability limits for the model's R^2 of 0.89 ($\rho = 0.9$) and 0.74 and ($\rho = 0.6$). Parameters for simultaneity ('S'): $\beta = 1$, $\alpha = 2$, $\sigma_\varepsilon^2 = 4$, $\sigma_\theta^2 = 1$. This yields probability limits for the model's R^2 of 0.84 ($\rho = 0.9$) and 0.81 ($\rho = 0.6$).

Table D.2: Simulation results: empirical size

	$n = 100$	$n = 500$	$n = 1000$	$n = 100$	$n = 500$	$n = 1,000$
	conditions (A.5) ('differences')			*conditions (A.4) ('generic')*		
$\rho = 0.9$		$T = 5$			$T = 5$	
ME	0.10	0.06	0.06	0.00	0.03	0.04
OV	0.10	0.06	0.06	0.00	0.02	0.03
S	0.12	0.05	0.05	0.00	0.03	0.05
$\rho = 0.6$						
ME	0.10	0.06	0.06	0.00	0.03	0.05
OV	0.06	0.05	0.06	0.00	0.02	0.03
S	0.10	0.05	0.05	0.00	0.03	0.05
$\rho = 0.9$		$T = 10$			$T = 10$	
ME	0.06	0.05	0.06	0.02	0.03	0.04
OV	0.07	0.05	0.05	0.02	0.05	0.05
S	0.07	0.05	0.06	0.03	0.05	0.05
$\rho = 0.6$						
ME	0.06	0.06	0.05	0.03	0.04	0.03
OV	0.06	0.05	0.05	0.02	0.05	0.06
S	0.07	0.05	0.05	0.02	0.04	0.05

Notes: In all cases, the simulation results are based on 1,000 simulation runs. The empirical size is obtained as the fraction of the number of simulation runs in which the test rejects the null hypothesis, while the null hypothesis is true. The simulated models correspond to three of the illustrative cases listed in Table 1 of the main text. Parameters for classical measurement error ('ME'): $\beta = 1$, $\sigma_\varepsilon^2 = 1$, $\sigma_\theta^2 = 1.44$, $\sigma_\eta^2 = 0$, $\delta = 0.3$. Parameters for omitted variables ('OV'): $\beta = 1$, $\gamma = 0$, $\sigma_\varepsilon^2 = 0.25$, $\sigma_\theta^2 = 0.36$, $\rho_{\theta\eta} = -0.6$. Parameters for simultaneity ('S'): $\beta = 1$, $\alpha = 0$, $\sigma_\varepsilon^2 = 4$, $\sigma_\theta^2 = 1$.

E. Call Report Data

We use year-end Call Report Data to create a balanced sample of U.S. commercial banks covering the 2011–2017 period. Filtering is done by excluding banks with total assets (in 2017 prices) less than USD 100 million, that are not part of a bank holding company, not located in a U.S. state, have no deposit insurance, or are not commercial according to their charter type. Furthermore, we remove bank-year observations with non-positive values of input factor expenses or total assets. We also delete bank-year observations with negative values for total loans and leases, total securities, or total equity. We leave out bank-year observations whose funding rate falls below the 1% sample quantile, whose wage rate exceeds the 99% sample quantile, or whose price of physical capital exceeds the 99% sample quantile. All level variables have been deflated using the Consumer Price Index for All Urban Consumers (CPIAUCSL) downloaded from the website of the Federal Reserve Bank of St. Louis. Table E.1 explains how the Call Report Data, downloaded from the FFIEC website, have been used to define the variables used in the empirical part of this study.

Table E.1: Definition of variables

Variable	Series/Definition
expenses on funding	RIAD4180+RIAD4170
salaries and employee benefits	RIAD4135
expenses of premises and fixed assets	RIAD4217
total funding	RCFD2200+RCFD2800
# of full-time equivalent employees on payrol	RIAD4150
premises and fixed assets	RCFD2145
funding rate	(expenses on funding)/(total funding)
wage rate	(salaries and employee benefits)/(# of full-time equivalent employees on payrol)
price of physical capital	(expenses of premises and fixed assets)/(premises and fixed assets)
total costs	sum of expenses on core deposits, purchased funds, labor and physical capital
total assets	RCFD2170
total loans and leases	RCFD1400
total securities	RCFD1754 + RCFD1773
off-balance sheet items	RIAD4079 – RIAD4080
total equity	RCFD3210
FDIC bank certificate ID	RSSD9050
date	RSSD9999
charter type	RSSD9048
physical state code	RSSD9210
primary insurer	RSSD9424
is bank part of BHC?	RSSD9364

Notes: This table explains how the variables in this study have been calculated from the data available in the Call Reports.

F. Additional estimation results

In this appendix, we apply the Wald test to two more elaborate specifications of the translog costs function than considered in the main text.

We start with a translog cost function with the same three input factors (funding, labor services, and physical capital) and single output factor (total assets) as before. Such a cost function has been widely used in the banking literature (e.g., Spierdijk and Zaouras, 2018; Fosu et al., 2017; Cubillas et al., 2017; Dong et al., 2016; Calderon and Schaeck, 2016; Carbó-Valverde et al., 2016). Using similar notation as in the main text, we specify

$$\begin{aligned}
\log(\widetilde{c}_{it}) &= \alpha_i + \gamma_t + \sum_{j=2}^{3}\beta_{j,p}\log(\widetilde{p}_{j,it}) + (1/2)\sum_{j=2}^{3}\beta_{j,pp}[\log(\widetilde{p}_{j,it})]^2 + \sum_{j=2}^{3}\sum_{k>j}\beta_{jk,pp}\log(\widetilde{p}_{j,it})\log(\widetilde{p}_{k,it}) \\
&+ \sum_{j=2}^{3}\beta_{j,pq}\log(\widetilde{p}_{j,it})\log(q_{it}) + \beta_q\log(q_{it}) + (1/2)\beta_{qq}[\log(q_{it})]^2 \\
&+ \beta_{eq}\log(EQ_{it}/q_{it}) + \varepsilon_{it}, \qquad \text{(F.1)}
\end{aligned}$$

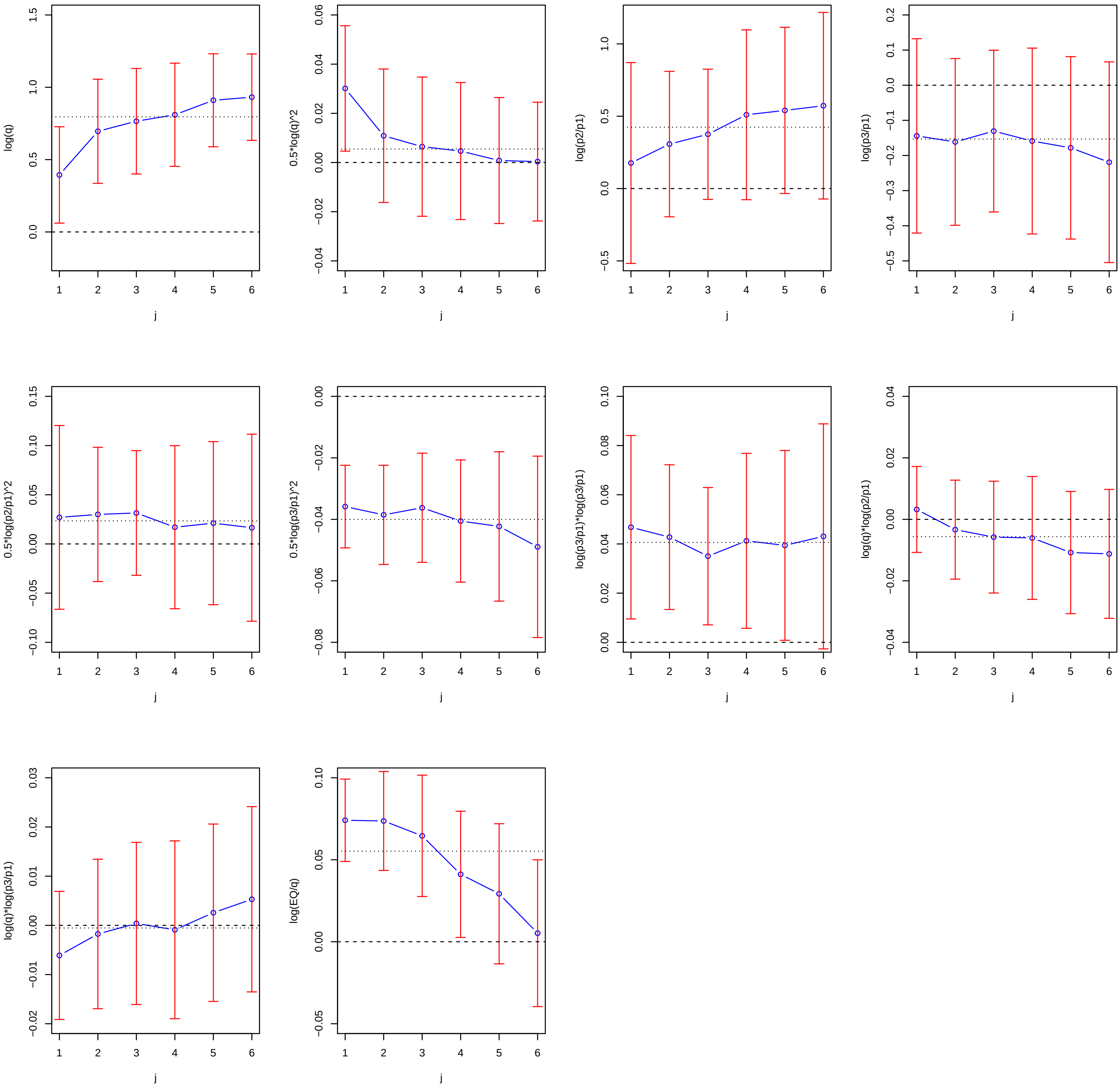
where EQ/q is the equity ratio. Throughout, variables with a tilde have been divided by the price of purchased funds $p_{1,it}$ to ensure linear homogeneity. For this model, the Wald test statistic equals 454.2 with p-value 0.0000 (degrees of freedom 10) and the resulting differences curves are visualized in Figure F.1.

We also consider a three-input three-output translog cost function, based on the same three inputs as before and three different outputs: total loans, total securities and off-balance sheet output. This choice of outputs is based on Wheelock and Wilson (2012) and results in a type of multi-product translog cost function that has been widely used in the banking literature (e.g., Tsionas et al., 2018; Spierdijk and Zaouras, 2018; Forssbæck and Shehzad, 2015; Kick and Prieto, 2015; Koetter et al., 2012). It is specified as

$$\begin{aligned}
\log(\widetilde{c}_{it}) &= \alpha_i + \gamma_t + \sum_{j=2}^{3}\beta_{j,p}\log(\widetilde{p}_{j,it}) + (1/2)\sum_{j=2}^{3}\beta_{j,pp}[\log(\widetilde{p}_{j,it})]^2 + \sum_{j=2}^{3}\sum_{k>j}\beta_{jk,pp}\log(\widetilde{p}_{j,it})\log(\widetilde{p}_{k,it}) \\
&+ \sum_{k=2}^{3}\sum_{\ell=1}^{3}\beta_{k\ell,pq}\log(\widetilde{p}_{k,it})\log(q_{\ell,it}) + \sum_{\ell=1}^{3}\beta_{\ell,q}\log(q_{\ell,it}) + (1/2)\sum_{\ell=1}^{3}\beta_{\ell,qq}[\log(q_{\ell,it})]^2 \\
&+ \sum_{\ell=1}^{3}\sum_{m>\ell}\beta_{\ell m,qq}\log(q_{\ell,it})\log(q_{m,it}) + \beta_{eq}\log(EQ_{it}/TA_{it}) + \varepsilon_{it}, \qquad \text{(F.2)}
\end{aligned}$$

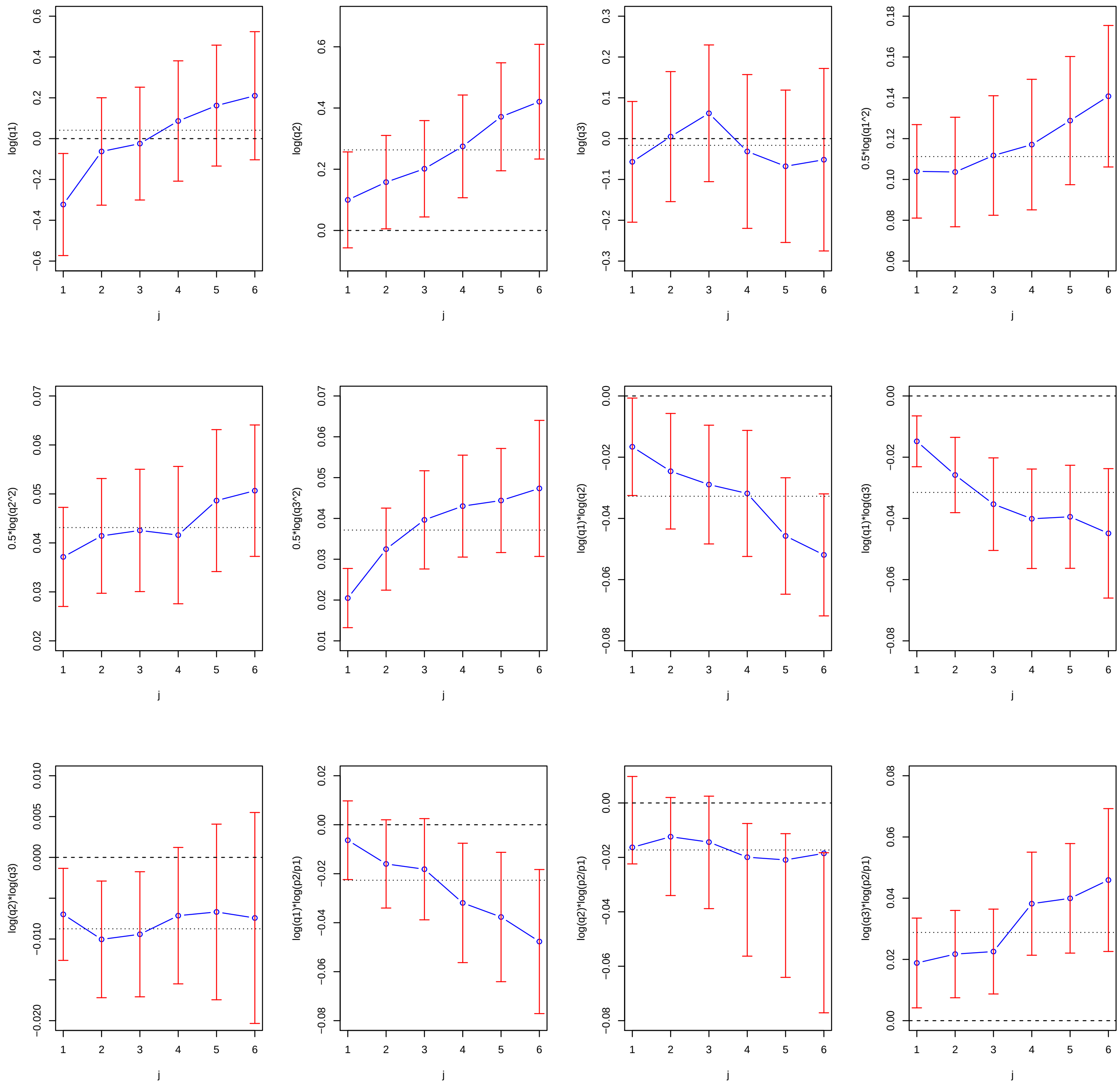
where $q_{1,it}$ denotes total loans, $q_{2,it}$ total securities and $q_{3,it}$ total assets of bank i in year t. As before, we use the equity ratio as a control variable, which is now denoted as EQ/TA. Appendix E provides the formal definitions of the input, output and control variables used in the cost function. For this model, the Wald test statistic equals 778.5 with p-value 0.0000 (degrees of freedom 21). The resulting differences curves are shown in Figures F.2 and F.3.

Figure F.1: Difference curves for the estimated coefficients of the cost function in (F.1)



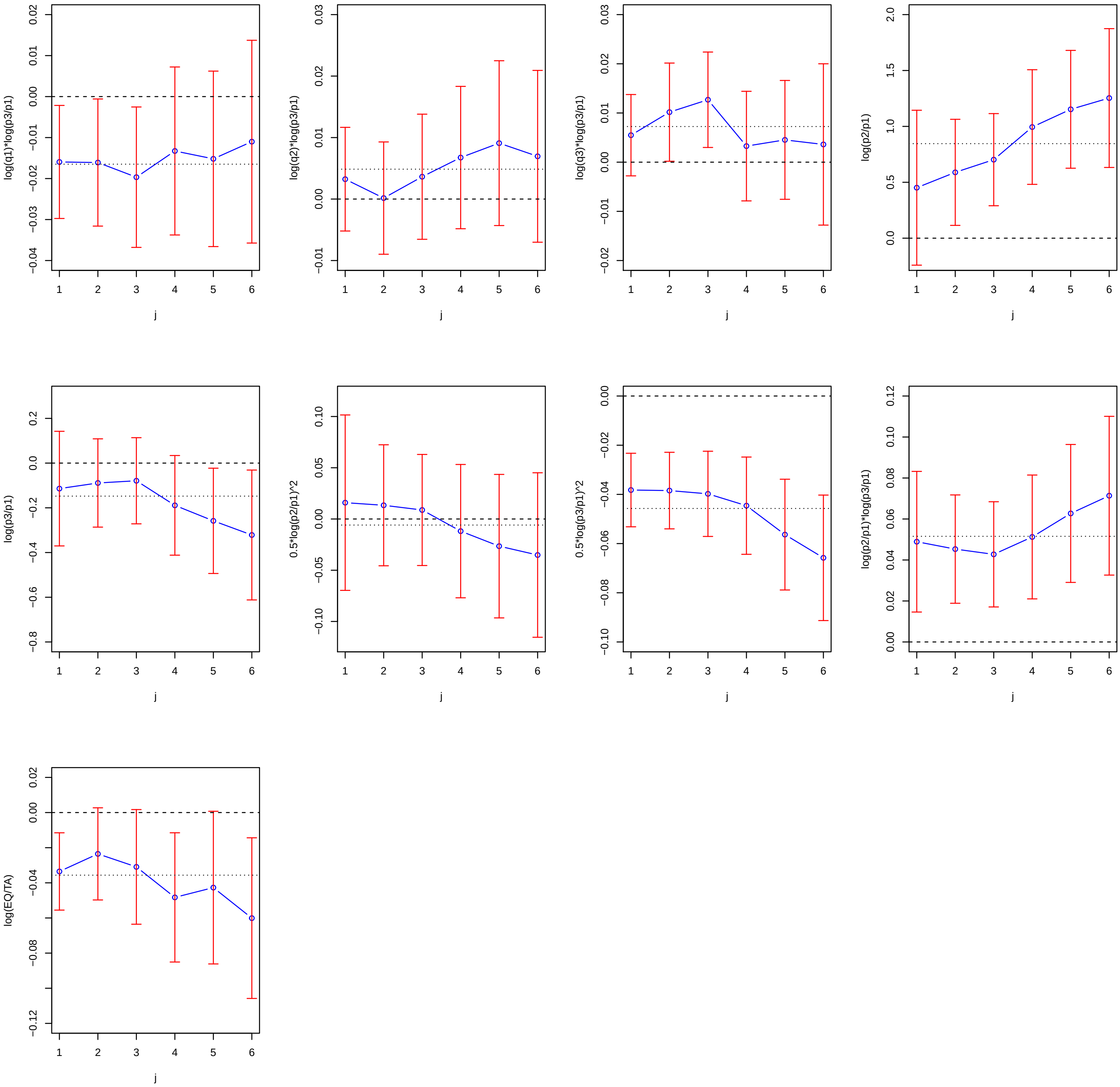
Notes: The intervals in red show the 95% asymptotic confidence interval for each point estimate based on cluster-robust standard errors.

Figure F.2: Difference curves for the estimated coefficients of the cost function in (F.2)



Notes: The intervals in red show the 95% asymptotic confidence interval for each point estimate based on cluster-robust standard errors.

Figure F.3: Patterns in the *j*-th difference estimates of the coefficients of the cost function in (F.2)



Notes: The intervals in red show the 95% asymptotic confidence interval for each point estimate based on cluster-robust standard errors.

G. Applications to existing data

This section considers two existing panel data sets from the literature, which each contain an explanatory variable that is suspected to be subject to measurement error. In the context of our theoretical results, these data sets provide a particularly relevant empirical case for our test.

G.1. Birth rates and welfare

Economic theory suggests that a government transfer program that reduces the cost of supporting a child should lead to a rise in birth rates. As pointed out by McKinnish (2008), childbearing is a commitment to current and future consumption. We may therefore expect fertility decisions to be relatively unresponsive to transitory fluctuations in welfare benefits. This would imply that welfare benefits are erroneous relative to the conceptual variable of interest, even though these benefits are generally reported without error in administrative records. As explained by Griliches and Hausman (1986), this kind of 'conceptual' measurement error is isomorphic to the errors-in-variables model with measurement error that is less persistent than the unobserved regressor and would render the within estimator inconsistent.

McKinnish (2008) aims to provide an empirical investigation of the presence of such conceptual measurement error in welfare benefits. She uses a panel data set consisting of U.S. state-level birth rates by white women in the age group 20–24.5 years and AFDC benefit levels for a family of four with no additional income. The panel data set with $n = 51$ and $T = 20$ covers the 1973–1992 period. The data set also contains a measure of the earnings per capita in each state. Both welfare benefits and earnings per capita are deflated and expressed in prices of the base year 1982–84.

We consider the linear panel regression model specified as

$$\log(y_{it}) = \alpha_i + \gamma_t + \beta_w \log(w_{it}) + \beta_e \log(e_{it}) + \varepsilon_{it}, \tag{G.1}$$

where y_{it} denotes the birth rate in state i in year t, α_i a state fixed effect, γ_t a year fixed effect, w_{it} the welfare benefit (i.e., the allegedly error-ridden regressor), and e_{it} the earnings per capita.

McKinnish (2008) estimates the linear panel regression model in (G.1) using data that is differenced over a time span of $j = 1, 3, 5, 7$ years. We denote the resulting coefficient estimates of the welfare benefit by $\hat{\beta}_{w,j}$. McKinnish (2008) compares the $\hat{\beta}_{w,j}$s for different values of j. In this way, she proceeds in a similar fashion as Goolsbee (2000). McKinnish (2008) establishes a monotonically increasing pattern in the $\hat{\beta}_{w,j}$s, which she contributes to the presence of conceptual measurement error.

Table G.1: Estimation results

	FE	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
McKinnish (2008)											
log(w)	0.123	0.020	0.059	0.074	0.087	0.094	0.110	0.116	0.123	0.130	0.133
	(0.021)	(0.377)	(0.139)	(0.142)	(0.112)	(0.104)	(0.066)	(0.045)	(0.023)	(0.013)	(0.014)
log(e)	0.120	0.150	0.328	0.373	0.344	0.301	0.237	0.186	0.191	0.141	0.113
	(0.194)	(0.227)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.011)	(0.034)	(0.110)	(0.172)
adj. R^2	0.044	0.010	0.079	0.107	0.097	0.085	0.070	0.056	0.061	0.051	0.049
Erickson and Whited (2000)											
q	0.018	0.016	0.019	0.019							
	(0.000)	(0.000)	(0.000)	(0.000)							
c	0.113	0.095	0.125	0.120							
	(0.011)	(0.020)	(0.007)	(0.043)							
cf	−0.099	−0.084	−0.104	−0.113							
	(0.030)	(0.050)	(0.029)	(0.065)							
adj. R^2	0.087	0.049	0.104	0.120							

Notes: This table shows the estimation results for the model in (G.1) applied to the data of McKinnish (2008) (upper panel) and for the model in (G.2) applied to the data of Erickson and Whited (2000) (lower panel). The column captioned 'FE' reports the estimation results based on the within estimator, while the columns captioned 'D_j' contain the estimates based on differences over time span j. The p-values associated with the estimated coefficients are in parentheses. The adjusted R^2s correspond to the transformed models whenever a panel data transformation has been used. Abbreviations used: w = welfare benefit; e = earnings per capita; c = cash flow per unit of capital; f = 0–1 variable for financially constraint firms (based on the availability of a bond rating). The time fixed effects are not reported, but have been taken out by using the dependent variable and the regressors in deviations from their means per time period.

We use the data set of McKinnish (2008) to formally test the within estimator's consistency. We estimate the linear panel regression model in (G.1) using data that is differenced over a time span of $j = 1, \ldots, 10$ years. The estimation results are summarized in the upper panel of Table G.1. This table also reports the FE estimation results.

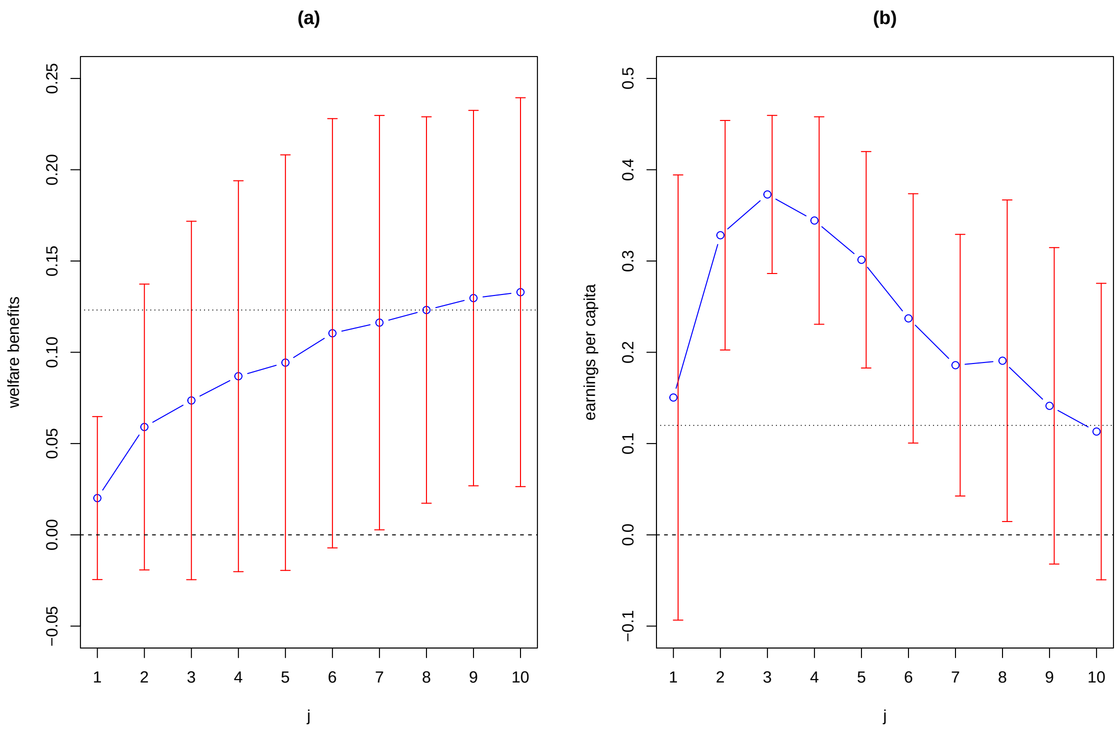
The Wald test rejects the null hypothesis $H_0 : \beta_{e,j} = \beta_{e,j+1}; \beta_{w,j} = \beta_{w,j+1}$ for $j = 1, \ldots, 9$ for each reasonable significance level; see Table G.2. Figures G.4(a) and (b) show the difference curves for the welfare-benefit and the earnings variables, which confirm the economic relevance of the rejection.

Table G.2: Test outcomes

	Wald	df	c.v.	p-value
McKinnish (2008)	139.6	18	28.87	0.00
Erickson and Whited (2000)	5.10	6	12.59	0.53

Notes: This table shows the results of the Wald tests for the model in (G.1) applied to the data of McKinnish (2008) (upper panel) and the model in (G.2) applied to the data of Erickson and Whited (2000) (lower panel). For each Wald test, we report the number of degrees of freedom ('df'), critical value ('c.v.') and p-value.

Figure G.4: Difference curves (McKinnish (2008) data)



Notes: This figure shows the difference curves for the estimated coefficients of the natural log of (a) the welfare benefit and (b) the earnings per capita. They correspond to the model in (G.1), applied to the data of McKinnish (2008). The intervals in red show the 95% asymptotic confidence interval for each point estimate. The two dashed lines indicate the zero line and the value of the within estimator.

Our test outcome suggests that McKinnish (2008) was right to suspect the within estimator of (G.1) to be inconsistent. Although the test outcome is consistent with the presence of measurement error in the welfare benefit variable, the type of misspecification remains an open question. For example, the data we used are aggregated across different cohorts and states that may respond differently to changes in welfare over time. Finding instrumental variables for the welfare benefit variable would be the first step in the empirical strategy that we outlined in Section 2.5.

G.2. Investments and Tobin's q

Erickson and Whited (2000) analyze the impact of Tobin's q on the investment rate. The theoretical motivation for studying this relation is the standard model of a perfectly competitive firm. This model is based on the maximization of net shareholder wealth, in the presence of convex adjustment costs following changes in the capital stock (e.g., Blundell et al., 1992). According to this model, Tobin's q has a positive effect on the investment rate. An empirical complication is the measurement-error problem associated with Tobin's q. This problem arises due to the difference between marginal q, the conceptual variable of interest, and measured q. Erickson and Whited (2000) discuss the possible sources of measurement error in measured q and propose an estimator that controls for such error by exploring higher-order moments. Their empirical analysis is based on a Compustat firm-level panel data set for the 1992–1995 period, with $n = 737$ and $T = 4$.

We consider the linear panel regression model given by

$$y_{it} = \alpha_i + \gamma_t + \beta_q q_{it} + \beta_c c_{it} + \beta_{fc} c_{it} f_i + \varepsilon_{it}, \tag{G.2}$$

where y_{it} denotes the ratio of investments to the replacement value of the capital stock for firm i in year t, α_i a firm fixed effect, γ_t a year fixed effect, q_{it} the proxy of marginal Tobin's q, c_{it} cash flow divided by the replacement value of the capital stock, and f_i a 0–1 variable indicating whether a firm is financially constrained or not. The indicator variable f_i is constructed on the basis of a firm's lack of bond rating and does not vary over time; its own marginal effect is therefore contained in the fixed effect α_i.

In the presence of measurement error in q, the within estimator of (G.2) will typically be inconsistent. We therefore use the data set of Erickson and Whited (2000) to estimate the linear panel regression model in (G.2) after differencing over a time span of $j = 1, 2, 3$ years.[2] Detailed estimation results are given in the lower panel of Table G.1. This table also reports the estimation results based on FE estimation.

The Wald test fails to reject the null hypothesis $H_0 : \beta_{q,j} = \beta_{q,j+1}, \beta_{c,j} = \beta_{c,j+1}; \beta_{fc,j} = \beta_{fc,j+1}$ for $j = 1, 2$; see again Table G.2. Although we do not reject, the difference curves for the covariates are provided in Figure G.5 for the sake of completeness.

How to go from here? Our test finds no evidence against the within estimator's inconsistency. This is a favorable outcome, but – as usual – we should remain aware of the possibility of that the test may have low power in certain cases. Low power could arise from limited data variability due to taking differences, yielding coefficient estimates with relatively large standard errors. In such a scenario, our test could fail to reject in the presence of misspecification. This explanation does not seem very likely in the present case, though. The strong significance of the estimated coefficients in the lower panel of Table G.1 suggests that the time-differenced data still contain a sufficient amount of variation.

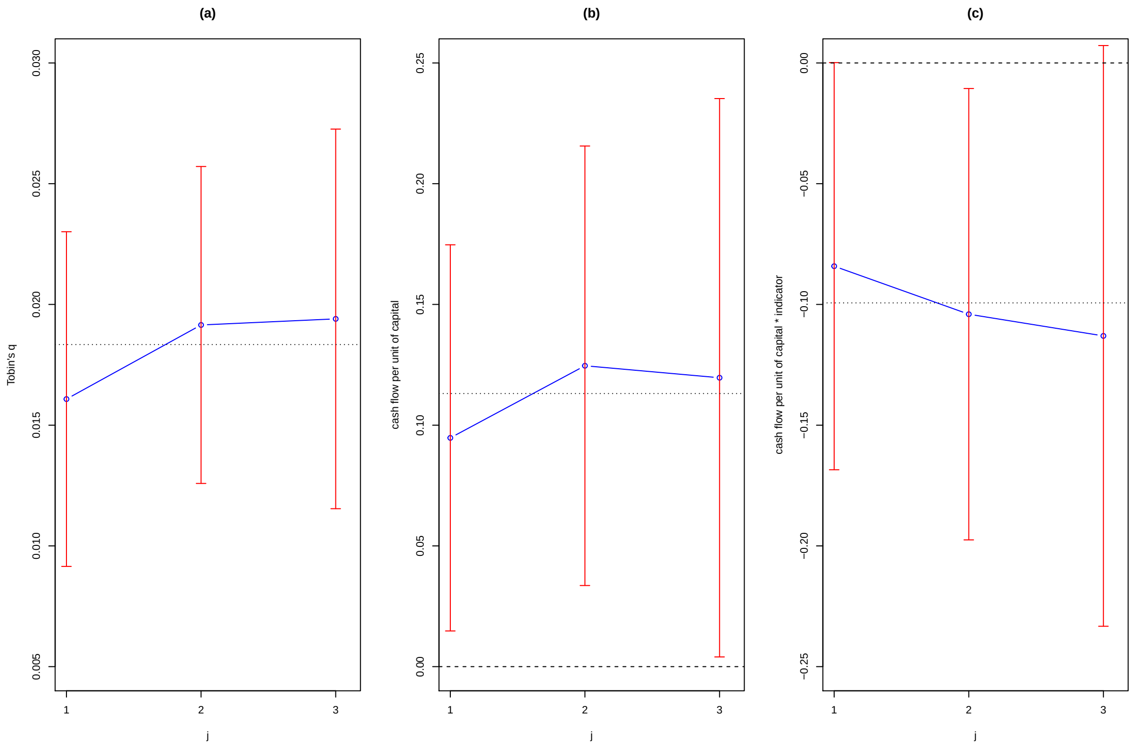
Another possibility is that the test outcome reflects the presence of two or more forms of endogeneity whose effects offset each other. To investigate this possibility in more detail, we could consider the potential endogeneity of the cash flow variable. For instance, cash flows may not only affect investments, but also vice versa. This would result in simultaneity and could offset the effects of measurement error in Tobin's q. Finding instrumental variables for Tobin's q and the cash-flow

[2] This data set of Erickson and Whited (2000) is available at http://toni.marginalq.com/publications.html.

variable would be the first step in the empirical strategy that we outlined in Section 2.5.

Given these considerations, it remains important to look for other evidence against the within estimator' consistency, such as coefficient signs that are unlikely from an economic perspective. Here, we find the coefficient signs that we would expect on the basis of economic theory: Tobin's q and the cash-flow variable both have a positive effect on the expected investment rate, which is smaller if firms are financially constraint.

Figure G.5: Difference curves (Erickson and Whited (2000) data)



Notes: This figure shows the difference curves for the estimated coefficients of (a) measured Tobin's q, (b) cash flows per unit of capital and (b) the interaction of cash flow per unit of capital and an indicator variable for financially constraint firms. They correspond to the model in (G.2), applied to the data of Erickson and Whited (2000). The intervals in red show the 95% asymptotic confidence interval for each point estimate based. The two dashed lines indicate the zero line and the value of the within estimator.

Acknowledgements

The authors thank Terra McKinnish for her willingness to share her data with us. The authors are also grateful to Vasilis Sarafidis and the other participants of the 25th International Panel Data Conference in Vilnius on July 4–5, 2019. Laura Spierdijk gratefully acknowledges financial support by a Vidi grant (452.11.007) in the 'Vernieuwingsimpuls' program of the Netherlands Organization for Scientific Research (NWO). Her work was also supported by the Netherlands Institute for Advanced Study in the Humanities and Social Sciences (NIAS-KNAW). The usual disclaimer applies.

References

Blundell, R., Bond, S., Devereux, M., Schiantarelli, F., 1992. Investment and Tobin's Q: evidence from company panel data. Journal of Econometrics 51, 233–257.

Calderon, C., Schaeck, K., 2016. The effects of government interventions in the financial sector on banking competition and the evolution of zombie banks. Journal of Financial and Quantitative Analysis 51, 1391 – 1436.

Cameron, A., Trivedi, P., 2005. Microeconometrics. Methods and Applications. Cambridge University Press, New York.

Carbó-Valverde, S., Rodríguez-Fernández, F., Udell, G., 2016. Trade credit, the financial crisis, and SME access to finance. Journal of Money, Credit and Banking 48, 113–143.

Cubillas, E., Fernández, A., González, F., 2017. How credible is a too-big-to-fail policy? International evidence from market discipline. Journal of Financial Intermediation 29, 46–67.

Dong, Y., Firth, M., Hou, W., Yang, W., 2016. Evaluating the performance of Chinese commercial banks: A comparative analysis of different types of banks. European Journal of Operational Research 252, 280–295.

Erickson, T., Whited, T., 2000. Measurement error and the relationship between investment and q,. Journal of Political Economy 108, 1027–1057.

Forssbæck, J., Shehzad, C., 2015. The conditional effects of market power on bank risk – cross-country evidence. Review of Finance 19, 1997–2038.

Fosu, S., Ntim, C., Coffie, W., Murinde., V., 2017. Bank opacity and risk-taking: Evidence from analysts' forecasts. Journal of Financial Stability 33, 81–95.

Goolsbee, A., 2000. The importance of measurement error in the cost of capital. National Tax Journal 53, 215–228.

Griliches, Z., Hausman, J., 1986. Errors in variables in panel data. Journal of Econometrics 31, 93–118.

Hall, A., 2005. Generalized method of moments. Oxford University Press.

Kick, T., Prieto, E., 2015. Bank risk and competition: Evidence from regional banking markets. Review of Finance 19, 1185–1222.

Koetter, M., Kolari, J., Spierdijk, L., 2012. Enjoying the quiet life under deregulation? Evidence from adjusted Lerner indices for US banks. Review of Economics and Statistics 94, 462–480.

McKinnish, T., 2008. Panel data models and transitory fluctuations in the explanatory variable, in: Fomby, T., Hill, R., Millimet, D., Smith, J., Vytlacil, E. (Eds.), Advances in Econometrics. volume 21: Modelling and Evaluating Treatment Effects in Econometrics, pp. 335–358.

Newey, W., 1983. Specification testing and estimation using a generalized method of moments. Ph.D. thesis. Massachusetts Institute of Technology.

Newey, W., 1985. Generalized method of moments specification testing. Journal of Econometrics 29, 229–256.

Spierdijk, L., Zaouras, M., 2018. Measuring banks market power in the presence of economies of scale: A scale-corrected Lerner index. Journal of Banking & Finance 87, 40–48.

Tsionas, E., Malikov, E., Kumbhakar, S., 2018. An internally consistent approach to the estimation of market power and cost efficiency with an application to U.S. banking. European Journal of Operational Research 270, 747–760.

Wansbeek, T., Meijer, E., 2000. Measurement error and latent variables in econometrics. North-Holland, Amsterdam.

Wheelock, D., Wilson, P., 2012. Do large banks have lower costs? New estimates of returns to scale for U.S. banks. Journal of Money, Credit and Banking 44, 171–199.

List of research reports

16001-GEM: Hoorn, A. van, How Are Migrant Employees Manages? An Integrated Analysis

16002-EEF: Soetevent, A.R., Te Bao, A.L. Schippers, A Commercial Gift for Charity

16003-GEM: Bouwmeerster, M.C., and J. Oosterhaven, Economic Impacts of Natural Gas Flow Disruptions

16004-MARK: Holtrop, N., J.E. Wieringa, M.J. Gijsenberg, and P. Stern, Competitive Reactions to Personal Selling: The Difference between Strategic and Tactical Actions

16005-EEF: Plantinga, A. and B. Scholtens, The Financial Impact of Divestment from Fossil Fuels

16006-GEM: Hoorn, A. van, Trust and Signals in Workplace Organization: Evidence from Job Autonomy Differentials between Immigrant Groups

16007-EEF: Willems, B. and G. Zwart, Regulatory Holidays and Optimal Network Expansion

16008-GEF: Hoorn, A. van, Reliability and Validity of the Happiness Approach to Measuring Preferences

16009-EEF: Hinloopen, J., and A.R. Soetevent, (Non-)Insurance Markets, Loss Size Manipulation and Competition: Experimental Evidence

16010-EEF: Bekker, P.A., A Generalized Dynamic Arbitrage Free Yield Model

16011-EEF: Mierau, J.A., and M. Mink, A Descriptive Model of Banking and Aggregate Demand

16012-EEF: Mulder, M. and B. Willems, Competition in Retail Electricity Markets: An Assessment of Ten Year Dutch Experience

16013-GEM: Rozite, K., D.J. Bezemer, and J.P.A.M. Jacobs, Towards a Financial Cycle for the US, 1873-2014

16014-EEF: Neuteleers, S., M. Mulder, and F. Hindriks, Assessing Fairness of Dynamic Grid Tariffs

16015-EEF: Soetevent, A.R., and T. Bružikas, Risk and Loss Aversion, Price Uncertainty and the Implications for Consumer Search

16016-HRM&OB: Meer, P.H. van der, and R. Wielers, Happiness, Unemployment and Self-esteem

16017-EEF: Mulder, M., and M. Pangan, Influence of Environmental Policy and Market Forces on Coal-fired Power Plants: Evidence on the Dutch Market over 2006-2014

16018-EEF: Zeng,Y., and M. Mulder, Exploring Interaction Effects of Climate Policies: A Model Analysis of the Power Market

16019-EEF: Ma, Yiqun, Demand Response Potential of Electricity End-users Facing Real Time Pricing

16020-GEM: Bezemer, D., and A. Samarina, Debt Shift, Financial Development and Income Inequality in Europe

16021-EEF: Elkhuizen, L, N. Hermes, and J. Jacobs, Financial Development, Financial Liberalization and Social Capital

16022-GEM: Gerritse, M., Does Trade Cause Institutional Change? Evidence from Countries South of the Suez Canal

16023-EEF: Rook, M., and M. Mulder, Implicit Premiums in Renewable-Energy Support Schemes

17001-EEF: Trinks, A., B. Scholtens, M. Mulder, and L. Dam, Divesting Fossil Fuels: The Implications for Investment Portfolios

17002-EEF: Angelini, V., and J.O. Mierau, Late-life Health Effects of Teenage Motherhood

17003-EEF: Jong-A-Pin, R., M. Laméris, and H. Garretsen, Political Preferences of (Un)happy Voters: Evidence Based on New Ideological Measures

17004-EEF: Jiang, X., N. Hermes, and A. Meesters, Financial Liberalization, the Institutional Environment and Bank Efficiency

17005-EEF: Kwaak, C. van der, Financial Fragility and Unconventional Central Bank Lending Operations

17006-EEF: Postelnicu, L. and N. Hermes, The Economic Value of Social Capital

17007-EEF: Ommeren, B.J.F. van, M.A. Allers, and M.H. Vellekoop, Choosing the Optimal Moment to Arrange a Loan

17008-EEF: Bekker, P.A., and K.E. Bouwman, A Unified Approach to Dynamic Mean-Variance Analysis in Discrete and Continuous Time

17009-EEF: Bekker, P.A., Interpretable Parsimonious Arbitrage-free Modeling of the Yield Curve

17010-GEM: Schasfoort, J., A. Godin, D. Bezemer, A. Caiani, and S. Kinsella, Monetary Policy Transmission in a Macroeconomic Agent-Based Model

17011-I&O: Bogt, H. ter, Accountability, Transparency and Control of Outsourced Public Sector Activities

17012-GEM: Bezemer, D., A. Samarina, and L. Zhang, The Shift in Bank Credit Allocation: New Data and New Findings

17013-EEF: Boer, W.I.J. de, R.H. Koning, and J.O. Mierau, Ex-ante and Ex-post Willingness-to-pay for Hosting a Major Cycling Event

17014-OPERA: Laan, N. van der, W. Romeijnders, and M.H. van der Vlerk, Higher-order Total Variation Bounds for Expectations of Periodic Functions and Simple Integer Recourse Approximations

17015-GEM: Oosterhaven, J., Key Sector Analysis: A Note on the Other Side of the Coin

17016-EEF: Romensen, G.J., A.R. Soetevent: Tailored Feedback and Worker Green Behavior: Field Evidence from Bus Drivers

17017-EEF: Trinks, A., G. Ibikunle, M. Mulder, and B. Scholtens, Greenhouse Gas Emissions Intensity and the Cost of Capital

17018-GEM: Qian, X. and A. Steiner, The Reinforcement Effect of International Reserves for Financial Stability

17019-GEM/EEF: Klasing, M.J. and P. Milionis, The International Epidemiological Transition and the Education Gender Gap

2018001-EEF: Keller, J.T., G.H. Kuper, and M. Mulder, Mergers of Gas Markets Areas and Competition amongst Transmission System Operators: Evidence on Booking Behaviour in the German Markets

2018002-EEF: Soetevent, A.R. and S. Adikyan, The Impact of Short-Term Goals on Long-Term Objectives: Evidence from Running Data

2018003-MARK: Gijsenberg, M.J. and P.C. Verhoef, Moving Forward: The Role of Marketing in Fostering Public Transport Usage

2018004-MARK: Gijsenberg, M.J. and V.R. Nijs, Advertising Timing: In-Phase or Out-of-Phase with Competitors?

2018005-EEF: Hulshof, D., C. Jepma, and M. Mulder, Performance of Markets for European Renewable Energy Certificates

2018006-EEF: Fosgaard, T.R., and A.R. Soetevent, Promises Undone: How Committed Pledges Impact Donations to Charity

2018007-EEF: Durán, N. and J.P. Elhorst, A Spatio-temporal-similarity and Common Factor Approach of Individual Housing Prices: The Impact of Many Small Earthquakes in the North of Netherlands

2018008-EEF: Hermes, N., and M. Hudon, Determinants of the Performance of Microfinance Institutions: A Systematic Review

2018009-EEF: Katz, M., and C. van der Kwaak, The Macroeconomic Effectiveness of Bank Bail-ins

2018010-OPERA: Prak, D., R.H. Teunter, M.Z. Babai, A.A. Syntetos, and J.E. Boylan, Forecasting and Inventory Control with Compound Poisson Demand Using Periodic Demand Data

2018011-EEF: Brock, B. de, Converting a Non-trivial Use Case into an SSD: An Exercise

2018012-EEF: Harvey, L.A., J.O. Mierau, and J. Rockey, Inequality in an Equal Society

2018013-OPERA: Romeijnders, W., and N. van der Laan, Inexact cutting planes for two-stage mixed-integer stochastic programs

2018014-EEF: Green, C.P., and S. Homroy, Bringing Connections Onboard: The Value of Political Influence

2018015-OPERA: Laan, N. van der, and W. Romeijnders, Generalized aplha-approximations for two-stage mixed-integer recourse models

2018016-GEM: Rozite, K., Financial and Real Integration between Mexico and the United States

2019001-EEF: Lugalla, I.M., J. Jacobs, and W. Westerman, Drivers of Women Entrepreneurs in Tourism in Tanzania: Capital, Goal Setting and Business Growth

2019002-EEF: Brock, E.O. de, On Incremental and Agile Development of (Information) Systems

2019003-OPERA: Laan, N. van der, R.H. Teunter, W. Romeijnders, and O.A. Kilic, The Data-driven Newsvendor Problem: Achieving On-target Service Levels.

2019004-EEF: Dijk, H., and J. Mierau, Mental Health over the Life Course: Evidence for a U-Shape?

2019005-EEF: Freriks, R.D., and J.O. Mierau, Heterogeneous Effects of School Resources on Child Mental Health Development: Evidence from the Netherlands.

2019006-OPERA: Broek, M.A.J. uit het, R.H. Teunter, B. de Jonge, J. Veldman, Joint Condition-based Maintenance and Condition-based Production Optimization.

2019007-OPERA: Broek, M.A.J. uit het, R.H. Teunter, B. de Jonge, J. Veldman, Joint Condition-based Maintenance and Load-sharing Optimization for Multi-unit Systems with Economic Dependency

2019008-EEF: Keller, J.T. G.H. Kuper, and M. Mulder, Competition under Regulation: Do Regulated Gas Transmission System Operators in Merged Markets Compete on Network Tariffs?

2019009-EEF: Hulshof, D. and M. Mulder, Renewable Energy Use as Environmental CSR Behavior and the Impact on Firm Profit

2019010-EEF: Boot, T., Confidence Regions for Averaging Estimators
2020001-OPERA: Foreest, N.D. van, and J. Wijngaard. On Proportionally Fair Solutions for the Divorced-Parents Problem

2020002-EEF: Niccodemi, G., R. Alessie, V. Angelini, J. Mierau, and T. Wansbeek. Refining Clustered Standard Errors with Few Clusters

2020003-I&O: Bogt, H. ter, Performance and other Accounting Information in the Public Sector: A Prominent Role in the Politicians' Control Tasks?

2020004-I&O: Fisch, C., M. Wyrwich, T.L. Nguyen, and J.H. Block, Historical Institutional Differences and Entrepreneurship: The Case of Socialist Legacy in Vietnam

2020005-I&O: Fritsch, M. and M. Wyrwich. Is Innovation (Increasingly) Concentrated in Large Cities? An Internatinal Comparison

2020006-GEM: Oosterhaven, J., Decomposing Economic Growth Decompositions.

2020007-I&O: Fritsch, M., M. Obschonka, F. Wahl, and M. Wyrwich. The Deep Imprint of Roman Sandals: Evidence of Long-lasting Effects of Roman Rule on Personality, Economic Performance, and Well-Being in Germany

2020008-EEF: Heijnen, P., On the Computation of Equilibrium in Discontinuous Economic Games

2020009-EEF: Romensen, G.J. and A.R. Soetevent, Improving Worker Productivity Through Tailored Performance Feedback: Field Experimental Evidence from Bus Drivers

2020010-EEF: Rao, Z., M. Groneck, and R. Alessie, Should I Stay or Should I Go? Intergenerational Transfers and Residential Choice. Evidence from China

2020011-EEF: Kwaak, C. van der, Unintended Consequences of Central Bank Lending in Financial Crises

2020012-EEF: Soetevent, A.R., Determinants choice set variation in demand estimation – with an application to the electric vehicle public charging market

2020013-EEF: Kwaak, C. van der, Old-Keynesianism in the New Keynesian model

2020014-EEF: Plaat, m. van der, Loan Sales and the Tyranny of Distance in U.S. Residential Mortgage Lending

2020015-I&O: Fritsch, M., and M. Wyrwich, Initial Conditions and Regional Performance in the Aftermath of Disruptive Shocks: The Case of East Germany after Socialism

2020016-OPERA: Laan, N. van der, and W. Romeijnders, A Converging Benders' Decomposition Algorithm for Two-stage Mixed-integer Recourse Models

2021001-OPERA: Baardman, L., K.J. Roodbergen, H.J. Carlo, and A.H. Schrotenboer, A Special Case of the Multiple Traveling Salesmen Problem in End-of-aisle Picking Systems

2021002-EEF: Wiese, R., and S. Eriksen, Willingness to Pay for Improved Public Education and Public Health Systems: The Role of Income Mobility Prospects.

2021003-EEF: Keller, J.T., G.H. Kuper, and M. Mulder, Challenging Natural Monopolies: Assessing Market Power of Gas Transmission System Operators for Cross-Border Capacity

2021004-EEF: Li, X., and M. Mulder, Value of Power-to-Gas as a Flexibililty Option in Integrated Electricity and Hydrogen Markets

2021005-GEM: Rozite, K., J.P.A.M. Jacobs, and D.J. Bezemer, Investor Sentiment and Business Investment

2021006-EEF: Spierdijk, L., and T. Wansbeek, Differencing as a Consistency Test for the Within Estimator www.rug.nl/feb