Archived version from NCDOCKS Institutional Repository http://libres.uncg.edu/ir/asu/

Appalachian
STATE UNIVERSITY
BOONE, NORTH CAROLINA

Symmetric Experimental Designs: Conditions For Equivalence Of Panel Data Estimators

By: Ronald L. Oaxaca & **David L. Dickinson**

Abstract

This paper specifies the panel data experimental design condition under which ordinary least squares, fixed effects, and random effects estimators yield identical estimates of treatment effects. This condition is relevant to the large body of laboratory experimental research that generates panel data. Although the point estimates and the true standard errors of the estimated average treatment effects are identical across the three estimators, the estimated standard errors differ. A standard F test as well as asymptotic reasoning guide the choice of which estimated standard errors are the appropriate ones to use for statistical inference.

Oaxaca, R. L. and **D. L. Dickinson** (2016). "Symmetric experimental designs: conditions for equivalence of panel data estimators." Journal of the Economic Science Association 2(1): 85-95. https://doi.org/10.1007/s40881-016-0022-x. Publisher version of record available at: https://link.springer.com/article/10.1007/s40881-016-0022-x

Symmetric experimental designs: conditions for equivalence of panel data estimators

Ronald L. Oaxaca[1,2,3,4] · David L. Dickinson[2,5,6]

Abstract This paper specifies the panel data experimental design condition under which ordinary least squares, fixed effects, and random effects estimators yield identical estimates of treatment effects. This condition is relevant to the large body of laboratory experimental research that generates panel data. Although the point estimates and the true standard errors of the estimated average treatment effects are identical across the three estimators, the estimated standard errors differ. A standard *F* test as well as asymptotic reasoning guide the choice of which estimated standard errors are the appropriate ones to use for statistical inference.

Keywords Experiments · Panel data · Econometrics · Estimators

JEL Classification C13 · C23 · C90

& Ronald L. Oaxaca
rlo@email.arizona.edu

1 Department of Economics, University of Arizona, Tucson, USA

2 IZA, Bonn, Germany

3 LISER, Esch-sur-Alzette, Luxembourg

4 PRESAGE, Texas, USA

5 Departmentof Economics, Appalachian State University, Boone, USA

6 ESI, Orange, USA

1 Introduction

In both experimental and non-experimental settings the advantages of panel data methods are widely recognized. Because of repeated observations in experiments, experimental data often constitute a panel. The presence of subject heterogeneity can lead to inefficient estimation by ordinary least squares (OLS).[1] Under these circumstances either fixed effects (FE) or random effects (RE) would be the estimator of choice. In this paper we demonstrate, both formally and with an empirical example from the literature, that for a panel data set generated by a symmetric design in which every subject faces every treatment exactly the same number of times, OLS, FE, and RE estimators yield identical treatment effect estimates: Result 1. Although the true standard errors would be identical for all three panel data estimators, it is shown below that the estimated standard errors for the treatment effects are identical between FE and RE but differ from those under OLS: Result 2. Typically, the experimentalist would want to use the FE estimated standard errors for statistical inference: Result 3.

In this paper we show how a common experimental design exactly conforms to the conceptual framework of Mundlak (1968). The importance of our results for researchers, especially experimental economists, is threefold. We show why the choice between the FE and the RE estimators is moot in important applied contexts, because these are one and the same estimator. Moreover, the estimated average treatment effects for FE and RE are identical to those obtained from (pooled) OLS. Lastly, we show that the only remaining choice is to decide whether to use the OLS standard errors or the FE/RE standard errors in finite samples. A standard *F* test as well as asymptotic reasoning guide the choice of which estimated standard errors are the appropriate ones to use for statistical inference.

2 Selected experimental studies

One can find a number of experimental studies that generate panel data which satisfy the symmetric treatment design. Dickinson et al. (2009) is an experimental exploration of the effects of alternative notions of employment risk on individual subject wage contracts. Data generated by a symmetric design were used to estimate average treatment effects. The institutional setting for this empirical example is open pit trading implemented by an oral double auction design. Smith et al. (1981) is a classic paper that implemented a computerized, double-auction design to document the potential for non-binding price floors and ceilings to bias price convergence relative to the competitive equilibrium. The symmetric design arose from each market group's exposure to the same treatments the same number of times. These symmetric design properties also apply at the level of individual buyer and seller behavior because individual subjects would constitute the cross-sectional observations for a study of the effects of the price control treatments on bids and offers, or surplus. In a recent paper Füllbrunn et al. (2015) conduct experiments to

[1] Often refered to as ''pooled'' OLS when applied to panel data.

examine the use of fall back options in second price auctions. A symmetric design is present because all subjects are exposed to the same treatments the same number of times.

In a recent paper, Dickinson et al. (2014) examine statistical discrimination in which two groups of subjects differentiated by their variance of productivity outcomes compete for jobs in an oral double auction setting. Since all subjects do not participate in all treatments, our estimator equivalence results do not apply to the full sample. Nevertheless, our equivalence results do apply to subsamples for which the subjects face the same exogenous treatments the same number of times. Cason et al. (2010) examines endogenous entry into a winner-take-all tournament versus a proportional-payment scheme. All subjects are exposed to the same treatments the same number of times, hence the symmetric design is present. Because binary choice RE probit models were estimated and RE probit is a nonlinear estimator, our results do not apply in this case. However, our estimator equivalance results would apply to panel data estimation of linear probability models of binary choice.

3 Experimental treatment models

Khuri (1992) illustrates the equivalence of FE and RE in a structural engineering experimental setting. In a somewhat different context Oaxaca et al. (2003) demonstrate the equivalence between pooled (cross-section and time-series) OLS estimates of the effects of time-invariant regressors and a two-stage feasible generalized least squares (FGLS) estimator of these effects. Mundlak's classic paper characterized as misguided the preoccupation of researchers with deciding whether fixed effects (FE) or random effects (RE) is the correct model. Mundlak shows that when one conditions on the sample individual (subject) effects, the FE and RE estimators are the same. Furthermore, Mundlak shows that in a correctly specified model, meaning that explicit account is taken of the relationship between the individual effects and the regressors, the RE estimator is identical to the FE estimator. In particular Mundlak shows that when individual effects are orthogonal to the regressors, the RE generalized least squares (GLS) estimator is identical to the FE within estimator.

We begin with a general specification of a balanced design experimental treatment model:

$$Y_{it} ¼ a þ X_{it}b þ \varepsilon_{it}$$

where Y_{it} is the experimental outcome variable, X_{it} is a $1 \times K$ vector of exogenous treatment indicator variables, b is a $K \times 1$ vector of average treatment effects, $i ¼ 1; \ldots; N$ (subjects) and $t ¼ 1; \ldots; T$: Counting one treatment indicator variable left out as the reference category, there are a total of K ? 1 treatments. Given the classical experimental design setup with exogenously determined treatments (fixed regressors), there is no correlation between X_{it} and the disturbance term ε_{it}: In the

case of the classical regression model where ε_{it} is i.i.d, OLS would be the estimator of choice.

The FE model arises when the intercept terms a_i are allowed to vary across subjects:

$$Y_{it} = a_i + X_{it}b + \varepsilon_{it}:$$

In this case the average treatment effects captured by the parameter vector b reflect the effects of each individual's treatment effect relative to the excluded treatment. The fixed effect parameter a_i can be viewed as an ''individual'' treatment effect since the model can written as

$$\begin{aligned} Y_{it} &= \sum_{k=1}^{K+1} d_{ki}X_{ikt} + \varepsilon_{it} \\ &= d_{1i} + \sum_{k=2}^{K+1} (d_{ki} - d_{1i})X_{ikt} + \varepsilon_{it}; \\ &= a_i + \sum_{k=2}^{K+1} b_k + \varepsilon_{it} \end{aligned}$$

where without loss of generality Treatment 1 is the left out reference category, $a_i = d_{1i}$; and $b_k = d_{ki} - d_{1i}\ \forall i$: Absent the usual left out reference group normalization or some restriction on the parameters, none of the individual treatment effects (d_{ki}) are identified. Subject to the usual excluded reference group normalization, we see that although the individual treatment effects a_i and d_{ki} differ across subjects, the differences between the individual treatment effects $d_{ki} - d_{1i}$ are invariant across subjects and are therefore the same as the average treatment effects. The FE model can be estimated by OLS with subject indicator variables, Least Squares Dummy Variable (LSDV), or equivalently in group deviation form (the within estimator).

Finally, the RE model arises if we assume that $a_i = a + u_i$:

$$Y_{it} = a + X_{it}b + \varepsilon_{it} + u_i;$$

where u_i is assumed to be i.i.d. and by the experimental design would be uncorrelated with X_{it}: Since the error process in the RE model yields a scalar variance/covariance matrix, i.e. constant variances and uncorrelated errors, the model is efficiently estimated by GLS (or FGLS when the true error variances have to be estimated).

4 Panel data estimators

An additonal estimator associated with panel data methods is the group means/between estimator. In this case the sample is collapsed down to a cross-section in which the variables are the sample means for each subject. The average treatment effect estimators corresponding to the OLS, FE, group means, and RE models are shown below (see Judge et al. 1980, p. 329–334).

$$\hat{\beta}^{ols} = X'\left(I_{NT} - \frac{i_{NT}i_{NT}'}{NT}\right)X \quad X'\left(I_{NT} - \frac{i_{NT}i_{NT}'}{NT}\right)Y$$

$$= (M_{xx})^{-1}M_{xy} \qquad (1)$$

$$\hat{\beta}^{fe} = \hat{\beta}^{w}$$

$$= X'\left(I_{NT} - I_N \otimes \frac{i_T i_T'}{T}\right)X \quad X'\left(I_{NT} - I_N \otimes \frac{i_T i_T'}{T}\right)Y$$

$$= (M^{w})^{-1}M^{w} \qquad (2)$$

$$\hat{\beta}^{b} = X'\left(I_N \otimes \frac{i_T i_T'}{T} - \frac{i_{NT}i_{NT}'}{NT}\right)X$$

$$\times X'\left(I_N \otimes \frac{i_T i_T'}{T} - \frac{i_{NT}i_{NT}'}{NT}\right)Y$$

$$= (M^{b})^{-1}M^{b} \qquad (3)$$

$$\hat{\beta}^{re} = r_\varepsilon^2\, wX'\left(I_N \otimes I_T - \frac{i_T i_T'}{T}\right)X + X'\left(I_N \otimes I_T - \frac{i_{NT}i_{NT}'}{NT}\right)X$$

$$+ \frac{1}{r_\varepsilon^2 + Tr_u^2}\, wX'\left(I_N \otimes \frac{i_T i_T'}{T} - \frac{i_{NT}i_{NT}'}{NT}\right)Y + X'\left(I_N \otimes I_T - \frac{i_T i_T'}{T}\right)Y$$

$$= (wM_{xx}^{b} + M_{xx}^{w})^{-1}\, wM_{xy}^{b} + M_{xy}^{w} \qquad (4)$$

where X is a $NT \times K$ observation matrix on the treatment indicator variables, Y is a $NT \times 1$ vector of observations on the experimental outcome variable, i_{NT} and i_T are $NT \times 1$ and $T \times 1$ vectors of 1's, M denotes the sample moment matrix, w denotes the within estimator, b denotes the between estimator, $w = \frac{r_\varepsilon^2}{r_\varepsilon^2 + Tr_u^2}$; r_ε^2 and r_u^2 are respectively the variances of ε_{it} and u_i:

Following Greene (2008, pp.191, 192, 202,203), we exploit the fact that the total

variation in the variables can be expressed as the sum of the within-variation and the between-variation:

$$M_{xx} ¼ M^{w}_{xx} þ M^{b}_{xx}$$
$$M_{xy} ¼ M^{w}_{xy} þ M^{b}_{xy}$$

Let p equal the number of rounds each treatment is administered. In the symmetric experimental design examined in this paper, each treatment appears pN times in the sample and $T ¼ pðK þ 1Þ$ is the number of observations per subject. Let T_{kit} be the indicator variable for the kth treatment corresponding to the ith subject in period t: Let

$$T_{ki} = \frac{\sum_{t=1}^{T} X_{kit}}{T} = \frac{p}{p(K+1)} = \frac{1}{(K+1)}$$

represent the sample mean of T_{kit} for the ith subject. Let

$$\bar{T}_k = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{kit}}{NT} = \frac{Np}{Np(K+1)} = \frac{1}{(K+1)}$$

represent the overall sample mean of T_{kit}. It is immediately clear that there is no between-variation because each observation involving the treatment observations is of the form

$$T_{ki} - \bar{T}_k = \frac{1}{(K+1)} - \frac{1}{(K+1)} = 0;$$

i.e. the average proportion of times each subject encounters each treatment is identical to the average proportion of times the treatment is encountered when averaged over all subjects and periods. Thus for example, in Dickinson et al. (2009) all subjects went through $K + 1 = 4$ treatments the same number of periods so that $T_{ki} = \bar{T}_k = 0.25$: What this means is that the cross-product matrices M^b_{xx} and M^b are null (zero) matrices. Consequently, $M_{xx} = M^w_{xx}$ and $M_{xy} = M^w_{xy}$. This establishes the result that $b^{ols} = b^w = b^{fe}$, i.e. the OLS and FE estimators of the treatment effects are identical. It remains to be shown that these treatment effects estimators are identical to the RE estimator in the symmetric experimental design.

In the case of the RE estimator expressed by (4), note that

$$X' \left(I_N \otimes \frac{i_T i_T'}{T} - \frac{i_{NT} i_{NT}'}{NT}\right) X = M^b_{xx}$$

and

$$X' \left(I_N \otimes I_T - \frac{i_T i_T'}{T}\right) Y = M^b_{xy}:$$

Since these cross-product matrices are null matrices, the RE estimator collapses to the within/FE estimator which is in turn the OLS estimator in this case:

$$\hat{b}^{re} = \left[X' \left(I_N \otimes I_T - \frac{i_T i_T'}{T}\right) X\right]^{-1} \left[X' \left(I_N \otimes I_T - \frac{i_T i_T'}{T}\right) Y\right]$$

$$\left(\mathbf{r} \quad \left(\quad \frac{\imath_T \imath'_0}{T} \right) \right) - \mathbf{r} \quad \left(\quad \frac{\imath_T \imath'_0}{T} \right) :$$

$$= X' \left(I_{NT} - I_N \otimes \frac{\imath_T \imath'_T}{T} \right) X \quad X' \left(I_{NT} - I_N \otimes \frac{\imath_T \imath'_T}{T} \right) Y$$

$$= \hat{\beta}_{fe} = \hat{\beta}_w = \hat{\beta}_{ols}$$

Thus, we have established *Result 1: the OLS, RE, and FE estimators are identical under our symmetric experimental design.*

We next consider estimation of the true variance/covariance matrix for the estimated average treatment effects given by

$$\mathrm{Var}(\hat{b}) = \sigma_\varepsilon^2 (M_{xx})^{-1}.$$

Although the estimators considered here are identical and use the same cross-product matrix M_{xx}; statistical software will use different formulas to estimate the residual variance σ_ε^2 depending on which estimator command is being used to estimate b. The question is then which variance/covariance matrix estimator formula, i.e. which estimate of σ_ε^2; should be used to estimate $\mathrm{Var}(\hat{b})$: Typically, the RE estimation strategy uses the FE estimates of σ_ε^2: Consequently, the error variance σ_ε^2 can be estimated from either the OLS residuals or from the FE residuals:

$$\hat{\sigma}^2_{e_{ols}} = \frac{e'_{ols} e_{ols}}{NT - (K+1)}$$

$$\hat{\sigma}^2_{e_{fe}} = \frac{e'_{fe} e_{fe}}{NT - (N+K)}.$$

This establishes *Result 2: the estimated standard errors for RE and FE are the same in our design but differ from those of OLS.*

Clearly the OLS and RE/FE error variance estimators are not independent. Note

$$e^{ols}_{it} - e^{fe}_{it} = (Y_{it} - \bar{Y}) - (X_{it} - \bar{X})\hat{b} - (Y_i - \bar{Y}) + (X_i - \bar{X})\hat{b}$$

$$= (Y_i - X_i b) - (Y - Xb)$$

$$= \hat{a}_i - \hat{a}$$

Since $\sum_i \sum_t e^{ols}_{it} = \sum_i \sum_t e^{fe}_{it} = 0$; it is easily seen that

$$\hat{a} = \hat{a}_i - \hat{a} = 0 \Rightarrow \hat{a} = \hat{a}_i = \frac{\sum_i \hat{a}_i}{N}$$: Therefore,

the estimated constant term in both the OLS and FE case is the average of the estimated individual fixed effects. We can express the OLS residuals in terms of the FE residuals:

$$e^{ols}_{it} = e^{fe}_{it} + \hat{a}_i - \hat{a}:$$

Squaring both sides of the preceding expression and summing over the t index yields:

$$\sum_t (e^{ols}_{it})^2 = \sum_t (e^{fe}_{it})^2 + T(\hat{a}_i - \hat{a})^2 + 2(\hat{a}_i - \hat{a}) \sum_t e^{fe}_{it}$$

$$\frac{1}{4}\left(\frac{\partial f_{it}^{e}}{\partial \theta_t}\right)\left(\frac{\partial K}{\partial \theta_t}\right)\left(\frac{\partial K}{\partial \theta_t}\right)^{2}.$$

Next, we sum over the index i to obtain

$$\sum_i \sum_t \left(e^{ols}_{it}\right)^2 \quad \sum_i \sum_t \left(e^{fe}_{it}\right)^2 \quad + p(K+1) \sum_i (\hat{a}_i - \hat{a})^2$$

or in vector notation

$$e^{ols\prime} e^{ols} = e^{fe\prime} e^{fe} + p(K+1) \sum_i (\hat{a}_i - \hat{a})^2 : \qquad (5)$$

If $a_i = a;\ 8\ i$, $\hat{s}^2_{e_{ols}}$ and $\hat{s}^2_{e_{fe}}$ are respectively unbiased and biased estimators of s^2. Likewise if $a_i \neq a$ for some i, $\hat{s}^2_{e_{ols}}$ and $\hat{s}^2_{e_{fe}}$ are respectively biased and unbiased estimators of s^2_E. Under the usual normality assumptions a finite sample test of OLS vs FE is based on the difference in restricted and unrestricted residuals:

$$\left(e^{ols\prime} e^{ols} - e^{fe\prime} e^{fe}\right) = (N-1) \sum_i (\hat{a}_i - \hat{a})^2 = (N-1) \cdot F_{N-1;NT-N-K-1}$$

$$\frac{e^{ols\prime} e^{ols} - e^{fe\prime} e^{fe}}{e^{fe\prime} e^{fe}/(NT-N-K-1)} = \frac{(N-1) \cdot F_{N-1;NT-N-K-1}}{e^{fe\prime} e^{fe}/(NT-N-K-1)}$$

Rejection of the hypothesis that the individual treatment effects relative to the left out treatment effect is constant across subjects, $a_i - a = 0\ 8i$, would favor the FE estimated standard errors.

Asymptotic considerations also provide guidance on the choice of which estimated standard errors to use. As is often the case with panel data methods, one needs to consider T asymptotics separately from N asymptotics and what this distinction means in a laboratory experimental setting. A little manipulation of (5) and noting that $p(K+1) = T$ yields

$$\hat{s}^2_{e_{ols}} = \hat{s}^2_{e_{fe}} \left[\frac{NT-(N+K)}{NT-(K+1)}\right] + \left[\frac{T \sum_i (\hat{a}_i - \hat{a})^2}{NT-(K+1)}\right] \qquad (6)$$

Under the i.i.d assumptions for E_{it}; letting $T \to \infty$ (which for fixed K means letting $p \to \infty$) and taking the probability limits of both sides of (6) establishes the result

$$\operatorname*{plim}_{(T\to\infty)} \hat{s}^2_{e_{ols}} = \operatorname*{plim}_{(T\to\infty)} \hat{s}^2_{e_{fe}} + \operatorname*{plim}_{(T\to\infty)} \hat{V}_a$$

$$= s^2_E + \operatorname*{plim}_{(T\to\infty)} \hat{V}_a$$

$$\sum_{i=1}^{N} \frac{(\hat{a} - \hat{a})^2}{N} \qquad \sum_{i=1}^{N} \frac{(a - a)^2}{N}$$

where $\hat{V}_a = \frac{\sum_{i=1}^{N} (\hat{a}_i - \bar{a})^2}{N}$ and $\operatorname{plim} \hat{V}_a = \frac{\sum_{i=1}^{N} (a_i - \bar{a})^2}{N}$; = 0: It can be shown

that as the number of experimental rounds $p \to \infty$; $\hat{\sigma}^2_{e_{fe}}$ is a consistent estimator of σ_E^2 whether or not $a_i - a = 0$ 8 i. On the other hand, it is evident that $\hat{\sigma}^2_{e_{ols}}$ is not a consistent estimator of ξ^2 if $a_i - a \neq 0$ for some i.

In an experimental setting it is natural to think of small (fixed) T. In this case if $N \to \infty$; we have the result

$$\operatorname*{plim}_{\delta N!1\Phi} \hat{\sigma}^2_{e_{\text{ols}}} \;\frac{1}{4}\; \operatorname*{plim}_{\delta N!1\Phi} \hat{\sigma}^2_{e_{\text{fe}}} \;\text{þ}\; \operatorname*{plim}_{\delta N!1\Phi} \hat{V}_a$$

$$\frac{1}{4}\, r^2_{\text{E}} \;\text{þ}\; \operatorname*{plim}_{\delta N!1\Phi} \hat{V}_a$$

$\delta N!1\Phi$ ^

where plim V_a ;= 0. It can be shown that as the number of experimental subjects N ! 1; $\hat{\sigma}^2_{e_{\text{fe}}}$ is a consistent estimator of r^2 whether or not $a_i - a$ ¼ 0 8i. As in the case with T asymptotics, $\hat{\sigma}^2_{e_{\text{ols}}}$ is not a consistent estimator of r^2_{E} if $a_i - a$ 6¼ 0 8i. Thus, plim $\hat{\sigma}^2_{e_{\text{ols}}}$;= plim $\hat{\sigma}^2_{e_{\text{fe}}}$ ¼ r^2_{E} in terms of N or T asymptotics. Therefore, unless the F test for OLS vs FE suggests otherwise, the FE estimated standard errors are the appropriate ones for statistical inference. Because it is usually the case that one can reject OLS against FE, we arrive at *Result 3: it is usually best to use the FE estimated standard errors for treatment effects.*

Another strategy might be to bootstrap the standard errors for $\hat{b}$: Because the estimator for the treatment effects in our framework is linear, bootstrapping may not elicit the sort of appeal it enjoys when applied to more complicated estimators. That being said, the fact that plim $\hat{\sigma}^2_{e_{\text{ols}}}$;= plim $\hat{\sigma}^2_{e_{\text{fe}}}$ ¼ r^2_{E} means one would probably want to exploit the assumed data generating process for the FE model to obtain the bootstrapped standard errors.

5 Discussion

Our basic point is a statistical one in an experimental setting but does not necessarily depend on the laboratory institution in any predictable way. While it is easiest to motivate our results in the context of individual choice experiments, the results would hold for individual behavior in market experiments and also when the market group is the unit of observations such as we showed would be possible in Smith et al. (1981). Other generalizations are possible such as cross-subject designs. A simple example is one in which one group of subjects is exposed only to Treatment 1, and another group of subjects is exposed only to Treatment 2. Suppose one wants to gauge the effects of Treatment 2 vs. Treatment 1 by comparing subjects across the two treatments. It can be shown that pooled OLS and RE are identical in terms of estimating the average treatment effect for 2 vs. 1. However, FE is not defined simply because every observation is associated with only Treatment 1 or Treatment 2 but not both. In this case one encounters perfect multicolinearity because the treatment indicators correspond exactly to the cross- section fixed effects. The subject demeaned values of the treatments are always zero.

All of the proceeding analysis assumes that there is no cross-sectional correlation in the panel data sets. Common shocks and other correlated factors are frequently

present in field data but could also be present with experimental data, depending on the experimental setting. How would the presence of cross-sectional dependence affect our results? De Hoyos and Sarafidis (2006) provides a comprehensive review of the literature on the cross-sectional dependence in panel data, including tests for

the presence of this phenomenon. If there is correlation between cross-section units, the equivalence of pooled OLS, FE, and RE still holds. Of course these are not fully efficient estimators and the associated estimated standard errors would not be correct. If the response is to merely fix up the standard errors with robust estimation of the variance/covariance matrix, the pooled OLS, FE and RE estimators would still be equivalent and inefficient but the asymptotic standard errors would be correct. If some sort of Full Information Maximum Likelihood approach is taken, then there is only one estimator of the model which is consistent and asymptotically efficient. The distinction between pooled OLS, FE and RE would be moot in this case.

6 Empirical example

In the symmetric experimental design of Dickinson et al. (2009), the focus was on the roles of various measures of risk on wage contracts. Consequently, only the RE estimates and associated standard errors were reported. In the table below we add the OLS and FE estimates for the employer sample. Four treatments are introduced that correspond to alternative measures of employment risk. The remaining conditioning variables consist of indicator variables for each of $p = 4$ rounds, and 4 indicator variables for the order in which a given treatment appears in a given session.

Wage contract determination ($N = 560$)

Variables	OLS		RE		FE	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Treatment 2	-0.003	0.034	-0.003	0.029	-0.003	0.029
Treatment 3	-0.028	0.041	-0.028	0.036	-0.028	0.036
Treatment 4	-0.062	0.033	-0.062	0.029	-0.062	0.029
Round 2	-0.119	0.032	-0.119	0.028	-0.119	0.028
Round 3	-0.143	0.032	-0.143	0.028	-0.143	0.028
Round 4	-0.149	0.032	-0.149	0.028	-0.149	0.028
Treatment order 2	-0.181	0.039	-0.181	0.034	-0.181	0.034
Treatment order 3	-0.183	0.033	-0.183	0.029	-0.183	0.029
Treatment order 4	-0.186	0.034	-0.186	0.029	-0.186	0.029
Constant	0.861	0.035	0.861	0.038	0.861	0.030

Data source: Dickinson et al. (2009)

The estimated treatment effects are seen to be identical between the three estimators while their estimated standard errors for OLS are different from the identical ones for RE/FE.

7 Concluding remarks

This paper shows that a particular type of experimental design is an example of the Mundlak (1968) equivalence between fixed effects and random effects estimators. In this special experimental setting, OLS, FE, and RE yield identical estimates of average treatment effects. Guidance is provided for estimating the residual variance to be used in constructing the appropriate standard errors.

Note that the equivalence of the OLS, FE, and RE panel data estimators is not altered by the addition of covariates that are orthogonal to the treatment effects. Examples of covariates that are orthogonal to the treatment indicators might be period/round dummies (as seen in the Sect. 6) or time-invariant regressors such as gender of the subject. Naturally, in the case of FE the estimated treatment effects are unaffected because the time-invariant regressors are not present due to being ''swept'' away by the FE estimator. Nevertheless, the estimated treatment effects also remain unchanged for OLS and RE. Moreover, the equivalence of these panel data estimators holds in the presence of non-orthogonal indicator variables whose values remain unchanged throughout the rounds for a given treatment but vary across treatments and subjects, e.g. treatment order indicator variables as seen in the empirical example. Of course the estimated treatment effects themselves will change with the addition of non-orthogonal covariates.

Acknowledgments We thank Todd Sørensen for his research assistance on an early version of this paper, and gratefully acknowledge the helpful comments and suggestions of Alfonso Flores-Lagunes, Tiemen Woutersen, Kei Hirano, the editor and two anonymous referees. All remaining errors are our own.

References

Cason, T. T., Masters, W. A., & Sheremeta, R. M. (2010). Entry into winner-take-all and proportional-prize contests: An experimental study. *Jounal of Public Economics*, *94*, 604–611.

De Hoyos, R. E., & Sarafidis, V. Testing for cross-sectional dependence in panel-data models. *The Stata Journal*, *6(4)*, 482–496.

Dickinson, D. L., & Oaxaca, R. L. (2009). Statistical discrimination in labor markets: An experimental analysis. *Southern Economic Journal*, *76(1)*, 16–31.

Dickinson, D. L., & Oaxaca, R. L. (2014). Wages, employment, and statistical discrimination: Evidence from the laboratory. *Economic Inquiry*, *52(4)*, 1380–1391.

Füllbrunn, S., Kreiner, S., & Palan, S. (2015). The value of a fallback option. *Central European Journal of Operations Research*, *23*, 375–388.

Greene, W. H. (2008). *Econometric Analysis*. Upper Saddle River: Pearson Prentice Hall.

Judge, G. G. et al. (1980). *The Theory and Practice of Econometrics*. New York: Wiley.

Khuri, A. I. (1992). Response surface models with random block effects. *Technometrics*, *34*(1), 26–37.

Mundlak, Y. (1968). On the pooling of time series and cross section data. *Econometrica*, *46(1)*, 69–85.

Oaxaca, R. L., & Geisler, I. (2003). Fixed effects models with time invariant variables: A theoretical note. *Economic Letters*, *80*(3), 373–377.

Smith, V. L., & Williams, A. W. (1981). On nonbinding price controls in a competitive market. *American Economic Review*, *71(3)*, 467—474.