Comparing the Bias of Dynamic Panel Estimators in Multilevel Panels: Individual versus Grouped Data

Nathan P. Hendricks and Aaron Smith

Abstract

We propose the Grouped Coefficients estimator to reduce the bias of dynamic panels that have a multilevel structure to the coefficient and factor loading heterogeneity. If groups are chosen such that the within-group heterogeneity is small, then the grouped coefficients estimator can lead to a substantial bias reduction compared to fixed effects and Arellano-Bond estimators. We also compare the magnitude of the bias of panel estimators with individual versus aggregate data and show that the magnitude of the bias also depends on the proportion of the heterogeneity that is within groups. In an application to estimating corn acreage response to price, we find that the grouped coefficients estimator gives reasonable results. Fixed effects and Arellano-Bond estimates of the coefficient on the lagged dependent variable appear to be severely biased with county-level data. In contrast, if we randomly assign the fields to groups and aggregate within the random groups, then pooled OLS of the randomly aggregated data gives a reasonable estimate of the coefficient on the lagged dependent variable.

1 Introduction

We study the estimation of dynamic regression models using panels of heterogeneous individuals or firms over time when the time series dimension of the panel (T) is small. This problem has been studied extensively in econometrics, but panels with a multilevel (i.e., cluster or hierarchical) structure have received little attention.[1] Multilevel structure is common in panel data, and we show that it can be exploited to reduce bias. Examples of multilevel structure include individuals who are more similar to others in their school, county, or state than those outside, and firms that are more similar to those in their sector or region than those outside.

Dynamic panels with heterogeneous intercepts and small T have received considerable attention in the literature. Nickell (1981) showed that fixed effects estimates are inconsistent for small T. Several articles have proposed consistent estimators that are based on using lags as instruments (e.g., Anderson and Hsiao, 1982; Arellano and Bond, 1991; Arellano and Bover, 1995) or deriving a bias correction (e.g., Kiviet, 1995; Hahn and Kuersteiner, 2002). The instrumental variables methods, however, are not consistent when either i) the coefficient on the lagged dependent variable or an autocorrelated regressor is heterogeneous (Pesaran and Smith, 1995) or ii) there is cross-sectional dependence (Sarafidis and Robertson, 2009).

Even when T is large, fixed effects estimates of dynamic panels with heterogeneous coefficients on the lagged dependent variable or an autocorrelated regressor are inconsistent (Robertson and Symons, 1992; Pesaran and Smith, 1995). In the case of large T, Pesaran and Smith (1995) propose the mean group estimator, where a separate regression is estimated for each individual and coefficients are averaged across individuals. Hsiao, Pesaran, and Tahmiscioglu (1999) find that the mean group estimator performs poorly when T is small and propose the use of a hierarchical Bayes estimator of the random coefficients.

[1] The typical panel formulation can be thought of as a multilevel structure with time and individuals as separate levels. Here we use the term "multilevel" to indicate multiple levels of heterogeneity in the cross-section.

There is also a literature that examines the impact of cross-section dependence on dynamic panel estimators with fixed T. Phillips and Sul (2007) show that incidental trends and cross-sectional dependence can substantially increase the bias of fixed effects in dynamic panels with heterogeneous intercepts and derive a bias correction. Sarafidis and Robertson (2009) show that time fixed effects can reduce the bias of generalized method of moments estimators with cross-sectional dependence as long as the variance of the factor loadings is small. Nauges and Thomas (2003) propose a double-differencing approach with generalized methods of moments to remove the influence of common factors.

We make two contributions to the dynamic panel data literature. First, we propose the use of a grouped coefficients estimator to reduce the bias in dynamic panels with i) heterogeneous coefficients, ii) cross-sectional dependence due to common factors or incidental trends, and iii) small T. Second, we investigate the difference in the bias of fixed effects under these conditions with individual versus aggregate data.

The grouped coefficients estimator exploits knowledge of the multilevel structure of the data to reduce the bias due to heterogeneous coefficients. The grouped coefficients estimator is implemented by estimating pooled OLS for groups of relatively homogeneous individuals, then an estimate of the mean marginal effect is obtained by averaging marginal effects across groups. If groups are defined such that there is no within-group coefficient heterogeneity and the groups are sufficiently large, then the grouped coefficients estimator consistently estimates the coefficients specific to each group and the overall mean coefficient. Estimating coefficients specific to each group removes the pooling bias from coefficient heterogeneity, while pooling within the groups reduces the bias of the coefficient on the lagged dependent variable due to small T.

A tradeoff exists in the specification of groups. If the groups are large, then pooling reduces the bias of the coefficient on the lagged dependent variable due to small T. But if groups are large, then the groups are likely to have more heterogeneity within groups and the pooling bias from heterogeneous coefficients is larger.

The grouped coefficients estimator also reduces the bias due to cross-sectional dependence when the factor loadings have a multilevel structure. If there is no within-group heterogeneity, then the bias from common factors is removed by including year fixed effects in the grouped coefficients estimator. The grouped coefficients estimator is especially beneficial when the overall variance of the factor loadings is large, but the variance of the factor loadings within groups is small.

Our work is related to studies that have used a grouped effects estimator in static models, where we use the term "grouped effects" to refer to an estimator with a separate intercept for each group. For example, Bester and Hansen (2009) consider a grouped effects estimator in nonlinear panel data models to reduce the bias from the incidental parameters problem. Deaton (1988) uses a grouped effects estimator with cross-section data.[2]

Our second contribution is to examine how the magnitude of the bias in standard panel estimators differs with individual versus aggregate data when there are heterogeneous coefficients and cross-sectional dependence in a multilevel panel. This is important because several studies use aggregate data (e.g., state or sector-level data). If most of the variation in parameters is within groups that are aggregated, then the aggregation reduces the parameter heterogeneity and panel estimators have less bias with aggregate data than individual data. However, if most of the parameter heterogeneity is between groups, then panel estimators typically have more bias with aggregate data than individual data. Furthermore, we show that even the direction of the bias of fixed effects may differ depending on whether individual or aggregate data are used.

We apply the grouped coefficients estimator to estimate the short-run and long-run corn acreage response to price in Iowa, Illinois, and Indiana. Our estimates with field-level data indicate that the long-run response to price is smaller than the short-run response to price. This result is expected given that corn and soybeans are commonly grown in a crop rotation in the region so that the probability of planting corn is lower if the field was previously

[2]Deaton (1988) and others refer to this estimator as the within-cluster estimator or cluster fixed effects.

planted to corn, *ceteris paribus*. Estimates of the dynamics of corn acreage response to price have important implications for understanding the effects of shocks to the corn market (e.g., ethanol production) on prices and the environment.

With field-level data, the coefficient on the lagged dependent variable is more negative with pooled fixed effects and the Arellano-Bond estimator than with the grouped coefficients estimator and pooled OLS gives a less negative coefficient. But all of the estimates with field-level data are between -0.70 and -0.54. For each estimator, using aggregate data increases substantially the coefficient on the lagged dependent variable. Estimates with county-level data are between -0.14 and 0.87, where the coefficient is positive for pooled OLS and pooled fixed effects. Consistent with our theory and simulations, our empirical estimates of the coefficient on the lagged dependent variable using county-level data are substantially larger than field-level estimates because the aggregation occurs over groups (counties) with large between-group heterogeneity. When we randomly assign fields to groups and aggregate the data within each of these groups—such that the aggregation should reduce the parameter heterogeneity—and estimate the model with pooled OLS, the coefficient on the lagged dependent variable is -0.52. Therefore, it is not the aggregation itself that results in the substantial bias of the coefficient on the lagged dependent variable, but the form of the aggregation (i.e., aggregating over groups with substantial between-group heterogeneity). The grouped coefficients estimates on prices do not differ substantially with fixed effects and Arellano-Bond estimates with county-level data. But we do find more variability in the coefficients on prices across different estimators with county rather than field-level data.

Our application to corn acreage response to price illustrates that the grouped coefficients estimator performs well when there is knowledge about the multilevel structure of the data and that caution is warranted when estimating dynamic panel models with aggregate data. In particular, the bias of standard dynamic panel estimators with aggregate data can far exceed the bias of pooled OLS with individual data when the aggregation occurs over groups with substantial between-group heterogeneity.

Section 2 defines a multilevel dynamic panel model, where a single common factor is a special case where all of the variation in the exogenous variable is between years, and describes the potential of the grouped coefficients estimator to reduce bias. Section 3 examines the magnitude of the bias of fixed effects with individual versus aggregate data. Section 4 describes our Monte Carlo simulations that compare the grouped coefficients estimator with pooled fixed effects and the Arellano-Bond estimator with individual and aggregate data, and section 4.2 presents results of the simulations. Section 5 presents our application to corn acreage response to price in Iowa, Illinois, and Indiana.

2 Exploiting the Multilevel Structure to Reduce Bias for Small T

We consider a heterogeneous dynamic panel with a common factor, where the coefficient heterogeneity has a multilevel structure. The heterogeneous dynamic panel that we consider is written as

$$
\begin{aligned}
y_{it} &= \gamma_i y_{i,t-1} + \beta_i x_{it} + \alpha_i + \varepsilon_{it}, \qquad (1)\\
x_{it} &= \mu_i(1-\rho) + \rho x_{i,t-1} + u_{it},\\
\gamma_i &\sim \mathrm{N}\left(\gamma, \sigma_\gamma^2\right), \qquad \beta_i \sim \mathrm{N}\left(\beta, \sigma_\beta^2\right), \qquad \kappa_i \sim \mathrm{N}\left(0, \sigma_\kappa^2\right),\\
\alpha_i &\sim \mathrm{N}\left(\alpha, \sigma_\alpha^2\right), \qquad \varepsilon_{it} \sim \mathrm{N}\left(0, \sigma_\varepsilon^2\right).
\end{aligned}
$$

We assume that x_{it} is strictly exogenous and is uncorrelated with the unobserved heterogeneity in the intercept, α_i. The independent variable has unconditional mean μ_i and an autocorrelation coefficient of ρ. The idiosyncratic errors, ε_{it}, are assumed to be independently and identically distributed. We assume that each individual can be assigned uniquely to a group. There are $G \ll N$ such groups, where N denotes the total number of individuals

in the sample. Denote by N_g the number of individuals in group g. The number of time periods is denoted T.[3]

A multilevel structure is given to the model by assuming that the parameter heterogeneity has a group-level component as well as an individual-level component,

$$\begin{aligned}
\gamma_i &= \gamma + \lambda_g + \lambda_i \\
&\quad \lambda_g \sim \mathrm{N}\left(0, (1-\delta)\,\sigma_\gamma^2\right), \qquad \lambda_i \sim \mathrm{N}\left(0, \delta\sigma_\gamma^2\right), \qquad E\left(\lambda_g \lambda_i\right) = 0 \\
\beta_i &= \beta + \eta_g + \eta_i \\
&\quad \eta_g \sim \mathrm{N}\left(0, (1-\delta)\,\sigma_\beta^2\right), \qquad \eta_i \sim \mathrm{N}\left(0, \delta\sigma_\beta^2\right), \qquad E\left(\eta_g \eta_i\right) = 0 \\
\alpha_i &= \alpha + \theta_g + \theta_i \\
&\quad \theta_g \sim \mathrm{N}\left(0, (1-\delta)\,\sigma_\alpha^2\right), \qquad \theta_i \sim \mathrm{N}\left(0, \delta\sigma_\alpha^2\right), \qquad E\left(\theta_g \theta_i\right) = 0,
\end{aligned} \tag{2}$$

where δ is the intraclass correlation coefficient. In other words, δ represents the proportion of the parameter variance that is within groups.

We also specify a multilevel structure for the variation of x_{it}. We assume that the variation of x_{it} over time has an aggregate component as well as an idiosyncratic component,

$$\begin{aligned}
u_{it} &= v_t + v_{it} \\
&\quad v_t \sim \mathrm{N}\left(0, (1-\phi)\,\sigma_u^2\right), \qquad v_{it} \sim \mathrm{N}\left(0, \phi\sigma_u^2\right),
\end{aligned} \tag{3}$$

where ϕ is the proportion of the variance of u_{it} that is due to idiosyncratic variation (i.e., within years). In many empirical applications, the variation in x_{it} is dominated by aggregate level changes over time. For example, our empirical application estimates the supply response of corn to price, where prices vary predominantly over time. Note that a common factor is

[3]Note that $\mathrm{N}(\cdot)$ denotes the normal distribution while N denotes the total number of individuals.

a special case of our model where all of the variation in x_{it} is between years ($\mu_i = \mu$, $\phi = 0$), and β_i is the factor loading.

Pesaran and Smith (1995) consider the case where the coefficients on the lagged dependent variable and the autocorrelated regressor are heterogeneous. If T is large, then the mean group estimator proposed by Pesaran and Smith (1995) gives consistent estimates of the short-run and long-run average marginal effects. The mean group estimator is the average of the coefficients estimated from separate least squares regressions for each individual. The mean group estimator is, however, inconsistent for small T since the coefficient on the lagged dependent variable is biased in the least squares regressions for each individual. Arellano-Bond type estimators that use lagged values as instruments are also inconsistent when the coefficients are heterogeneous (Pesaran and Smith, 1995). Hsiao, Pesaran, and Tahmiscioglu (1999) perform Monte Carlo simulations and find that the mean group estimator performs poorly when T is small relative to N. Instead, Hsiao, Pesaran, and Tahmiscioglu (1999) propose the use of a hierarchical Bayes estimator of the random coefficients and find that it performs well in Monte Carlo simulations.

Sarafidis and Robertson (2009) consider the effect of common factors on generalized method of moments estimates under the case of a homogenous coefficient on the lagged dependent variable. They show that generalized method of moments estimates can be severely biased in the presence of the common factor, but that including year fixed effects reduces the bias. The benefit of including year fixed effects is reduced as the variance of the factor loadings increases. Intuitively, when the factor loadings are homogeneous, a model with year fixed effects is correctly specified.

We propose a grouped coefficients estimator that exploits knowledge of the multilevel structure of the data to reduce the bias in a heterogeneous dynamic panel that may include a common factor.[4] To implement the grouped coefficients estimator, the researcher defines

[4]We do not refer to the grouped coefficients estimator as a mean group estimator to avoid confusion with the estimator proposed by Pesaran and Smith (1995) where regressions are estimated separately for each individual.

groups of relatively homogeneous individuals and estimates pooled OLS regressions for each group. The grouped coefficients estimate of the mean coefficient is the weighted average of the coefficients across the groups, where the weights correspond to the size of each group. The bias from an omitted common factor is reduced by including year fixed effects in the group-specific regressions. Alternatively, if the model has incidental trends as in Phillips and Sul (2007), then the bias is reduced by adding a trend to the group-specific regressions.

It is straightforward to see that if groups are defined such that there is no within-group coefficient heterogeneity and the groups are sufficiently large, then the grouped coefficients estimator consistently estimates the coefficients specific to each group. In this case, estimating coefficients specific to each group removes the pooling bias from coefficient heterogeneity, and pooling within groups reduces the bias of the coefficient on the lagged dependent variable due to fixed T.

An advantage of the grouped coefficients estimator over the Bayesian random coefficients estimator is that the grouped coefficients estimator does not impose a specific distributional form on the coefficient heterogeneity and allows any form of correlation between the coefficients at the group level. The grouped coefficients estimator also allows the unobserved group-specific heterogeneity to be correlated with the regressors. Furthermore, the grouped coefficients estimator is appealing because it is computationally simple compared to Bayesian estimators. For example, in our application to corn acreage in section 5, we use a dataset with roughly 1 million observations per year.

Bester and Hansen (2009) consider a grouped effects estimator in nonlinear panel data models where the intercept varies across individuals. Nonlinear panel estimators with fixed effects suffer from the incidental parameters problem. Bester and Hansen (2009) find a tradeoff from the bias of incidental parameters and the bias from misspecification of the unobserved heterogeneity. Estimating group-specific intercepts reduces the bias from incidental parameters if groups are specified with a large number of individuals. But if groups have a

large number of individuals, then there is also likely to be more unobserved heterogeneity within groups.

An analogous tradeoff exists in our specification of groups in a multilevel dynamic model.[5] If the groups are large, then pooling minimizes the bias of the coefficient on the lagged dependent variable due to fixed T. But if groups are large, then the groups are likely to have more within-group heterogeneity and the pooling bias from heterogeneous coefficients is larger.

An alternative approach to reduce the bias of multilevel dynamic panel estimates is to aggregate the data by groups, where the groups are defined such that all of the parameter heterogeneity is within groups (see the discussion below in sections 3.1 and 3.2). For example, this approach could be implemented by randomly assigning individuals to groups. Intuitively, aggregation can eliminate the bias if the parameter heterogeneity is removed in the aggregation.

The aggregation approach, however, is problematic because aggregation of a heterogeneous dynamic panel leads to a different dynamic structure at the aggregate level than at the individual level (e.g., see Granger, 1980). Lewbel (1994) shows that the coefficient on the first lagged dependent variable of an aggregate time series is the mean of the individual coefficients on the lagged dependent variable under the assumption that the coefficient on the lagged dependent variable is uncorrelated with the coefficient on the regressor. Under more general assumptions, an aggregate model with a single lagged dependent variable gives inconsistent estimates of the average coefficients (Pesaran and Smith, 1995).

[5]The incidental parameters problem underlies the bias of fixed effects in linear dynamic panels as well (see Lancaster, 2000).

3 Bias of Fixed Effects with Individual versus Aggregate Data

Several studies use aggregate data rather than individual data to estimate a dynamic model. In this section, we compare the bias of pooled fixed effects with individual versus aggregate data and show how the bias depends on whether the data are aggregated over relatively homogenous or relatively heterogeneous individuals. To keep results tractable, we consider the bias of pooled fixed effects with individual versus aggregate data in two cases: the coefficient on an autocorrelated regressor is heterogeneous (section 3.1), and the coefficient on the lagged dependent variable is heterogeneous (section 3.2).

3.1 Heterogeneous Coefficient on the Autocorrelated Regressor

First, we consider a dynamic model with a heterogeneous coefficient on an autocorrelated regressor,

$$y_{it} = \gamma y_{i,t-1} + \beta_i x_{it} + \alpha_i + \varepsilon_{it}, \tag{4}$$
$$x_{it} = \mu_i(1-\rho) + \rho x_{i,t-1} + u_{it}.$$

Pooled fixed effects gives biased estimates of the average coefficients in (4), even when $N \to \infty$ and $T \to \infty$ (Pesaran and Smith, 1995). The source of the bias in pooled fixed effects in this case is the serial correlation of the composite error term, which is therefore correlated with $y_{i,t-1}$.

Although we are primarily concerned with the case of small T in this paper, it is instructive to examine the bias of fixed effects as $T \to \infty$ since this will isolate the bias that is due to the heterogeneous coefficient (β_i). With fixed T there is also the well-known negative bias of the coefficient on the lagged dependent variable in fixed effects due to the differencing.

We denote the fixed effects estimates with individual data as $\hat{\gamma}^{ind}$ and $\hat{\beta}^{ind}$. Pesaran and Smith (1995) show that the bias of pooled fixed effects when $N \to \infty$ and $T \to \infty$ is

$$\underset{N\to\infty,T\to\infty}{plim}\left(\hat{\gamma}^{ind}\right) - \gamma = \frac{\rho\left(1-\gamma\rho\right)\left(1-\gamma^2\right)\sigma_\beta^2}{\Psi_1}, \tag{5}$$

$$\underset{N\to\infty,T\to\infty}{plim}\left(\hat{\beta}^{ind}\right) - \beta = -\frac{\beta\rho^2\left(1-\gamma^2\right)\sigma_\beta^2}{\Psi_1}, \tag{6}$$

where

$$\Psi_1 = \frac{\sigma_\varepsilon^2}{\sigma_u^2}\left(1-\rho^2\right)\left(1-\gamma\rho\right)^2 + \left(1-\gamma^2\rho^2\right)\sigma_\beta^2 + \left(1-\rho^2\right)\beta. \tag{7}$$

The estimate of γ is biased up and the estimate of β is biased down. The bias is increasing in the variance of β_i, and decreasing in the ratio of variances ${}^{\sigma_\varepsilon^2}/_{\sigma_u^2}$.

Next, consider the bias of fixed effects when aggregate data are used instead of individual data. Granger (1980) shows that aggregating a heterogeneous dynamic panel can lead to a substantially different dynamic structure at the aggregate level. In equation (4), however, the coefficient on the lagged dependent variable is constant, so estimating the aggregate dynamics with a single lagged dependent variable poses no particular problem.

Aggregating to the group level affects the variance of the parameters; and thus, change the magnitude of the bias of fixed effects. Let the average group-level coefficient on the independent variable be defined by $\bar{\beta}_g = \frac{1}{N_g}\sum_{i\in g}\left(\beta + \eta_g + \eta_i\right)$. The variance of the group-level coefficients is $Var\left(\bar{\beta}_g\right) = Var\left(\eta_g\right) + \frac{1}{N_g}Var\left(\eta_i\right)$ or $Var\left(\bar{\beta}_g\right) = \left(1-\delta\right)\sigma_\beta^2 + \frac{\delta\sigma_\beta^2}{N_g}$. Aggregation also reduces the variation in x_{it} that is relevant for the bias of fixed effects, $Var\left(\bar{u}_{gt}\right) = \left(1-\phi\right)\sigma_u^2 + \frac{\phi\sigma_u^2}{N_g}$. The variance of the idiosyncratic errors at the group level is $Var\left(\bar{\varepsilon}_{gt}\right) = \frac{\sigma_\varepsilon^2}{N_g}$.

Substituting the variance of the parameters with aggregate data into equations (5) and (6), the bias of pooled fixed effects with aggregate data is

$$\underset{N\to\infty,T\to\infty}{plim}\left(\hat{\gamma}^{grp}\right)-\gamma=\frac{\rho\left(1-\gamma\rho\right)\left(1-\gamma^{2}\right)\left(\left(1-\delta\right)\sigma_{\beta}^{2}+\frac{\delta\sigma_{\beta}^{2}}{N_{g}}\right)}{\Psi_{2}}, \tag{8}$$

$$\underset{N\to\infty,T\to\infty}{plim}\left(\hat{\beta}^{grp}\right)-\beta=-\frac{\beta\rho^{2}\left(1-\gamma^{2}\right)\left(\left(1-\delta\right)\sigma_{\beta}^{2}+\frac{\delta\sigma_{\beta}^{2}}{N_{g}}\right)}{\Psi_{2}}, \tag{9}$$

where

$$\Psi_{2}=\frac{\sigma_{\varepsilon}^{2}/N_{g}}{\left(1-\phi\right)\sigma_{u}^{2}+\frac{\phi\sigma_{u}^{2}}{N_{g}}}\left(1-\rho^{2}\right)\left(1-\gamma\rho\right)^{2}+\left(1-\gamma^{2}\rho^{2}\right)\left(\left(1-\delta\right)\sigma_{\beta}^{2}+\frac{\delta\sigma_{\beta}^{2}}{N_{g}}\right)+\left(1-\rho^{2}\right)\beta. \tag{10}$$

Aggregating to the group-level decreases the variance of σ_{β}^{2}, which decreases the bias of fixed effects. On the other hand, aggregating to the group-level also decreases the variance of the idiosyncratic error relative to the time-varying portion of the variance of x_{it}, $\left(\frac{\sigma_{\varepsilon}^{2}/N_{g}}{(1-\phi)\sigma_{u}^{2}+\frac{\phi\sigma_{u}^{2}}{N_{g}}}\leq\frac{\sigma_{\varepsilon}^{2}}{\sigma_{u}^{2}}\right)$, which increases the bias of fixed effects.

To examine whether the bias of fixed effects is larger with individual or aggregate data, consider two special cases where all of the parameter variation is within groups and when all of the parameter variation is between groups. In the case where all the parameter variation is within groups (i.e., $\delta=1$), then using aggregate data substantially reduces the parameter heterogeneity and the fixed effects estimator converges to the true parameter values as the number of individuals in each group increases. When all of the parameter variation is between groups (i.e., $\delta=0$), then the parameter variation does not decrease by using aggregate data. But the variance of the idiosyncratic error relative to the time-varying portion of the variance of x_{it} decreases, and the bias of pooled fixed effects is larger with aggregate data than with individual data. In summary, using aggregate data decreases the bias of fixed effects when most of the parameter variation is within groups. However, using aggregate data can increase

the bias of fixed effects when a portion of the parameter heterogeneity is between groups and a portion of the variation in x_{it} is common to all individuals.

If x_{it} is a common factor, then none of the variation is idiosyncratic ($\phi = 0$). Since a common factor is likely to be autocorrelated—as we would think is true for macro-economic shocks—then the coefficient on the lagged dependent variable captures the autocorrelation in the unobserved common factors in addition to the dynamics of the model. Even if individual and year fixed effects are included with large T, the common factor causes the coefficient on the lagged dependent variable to be biased since individuals with a greater response to the common factor are also likely to have a larger lagged dependent variable.

Similar to our previous arguments, a dynamic regression with individual and year fixed effects has a larger bias with aggregate data than individual data if most of the heterogeneity of the factor loadings is between groups since the aggregation does not reduce the heterogeneity of the factor loadings and the variance of the common factor is a larger proportion of the variation in the dependent variable. This result is related to Granger (1987), who demonstrated that common factors may have little explanatory power at the individual level, but substantial explanatory power at the aggregate level. Omitting year fixed effects to control for common factors causes more bias with aggregate data than individual data.

While we only consider the case of a single common factor, Phillips and Sul (2007) note that for individual data the number of factors will typically be large and the number of time periods small. They find that cross-section dependence has little effect on the dynamic panel estimators when factors are independent and the number of factors is large. In contrast, aggregate panels will have a small number of common factors that increases the bias of panel estimators (Phillips and Sul, 2007).

3.2 Heterogeneous Coefficient on the Lagged Dependent Variable

Next, we consider a model where the coefficient on the lagged dependent variable is heterogeneous, but the coefficient on the independent variable is homogeneous and there is no autocorrelation in the independent variable,

$$y_{it} = \gamma_i y_{i,t-1} + \beta x_{it} + \alpha_i + \varepsilon_{it}, \tag{11}$$
$$x_{it} = \mu_i + u_{it}.$$

Using equation (A.20)-(A.25) in Pesaran and Smith (1995), we obtain the following formulas for the bias of pooled fixed effects with individual data

$$\underset{N\to\infty,T\to\infty}{plim}\left(\hat{\gamma}^{ind}\right) - \gamma = \frac{\frac{\sigma_\varepsilon^2}{\sigma_u^2} E\left(\frac{\lambda_g+\lambda_i}{1-\gamma_i^2}\right) + \beta^2 \sum_{j=0}^{\infty} E\left[\left(\lambda_g+\lambda_i\right)\gamma_i^{2j}\right]}{\Psi_3}, \tag{12}$$

$$\underset{N\to\infty,T\to\infty}{plim}\left(\hat{\beta}^{ind}\right) - \beta = 0, \tag{13}$$

where $\gamma_i = \bar{\gamma} + \lambda_g + \lambda_i$ and

$$\Psi_3 = \frac{\sigma_\varepsilon^2}{\sigma_u^2} E\left(\frac{1}{1-\gamma_i^2}\right) + \beta^2 \sum_{j=0}^{\infty} E\left(\gamma_i^{2j}\right). \tag{14}$$

The fixed effects estimate of the coefficient on the independent variables is not biased in this case, but the lack of autocorrelation in x_{it} is a necessary assumption for the bias to disappear. The sign of the bias in $\hat{\gamma}^{ind}$ is indeterminate, but is likely to depend on the sign of γ_i.

Note that the magnitude of the bias in equation (12) depends on the the variance of the idiosyncratic error relative to the time-varying portion of the variance of x_{it}. Therefore,

using aggregate data instead of individual data affects the magnitude of the bias of fixed effects, but it is not clear whether the aggregation increases or decreases the bias.

Also note that if all of the parameter variation is within groups, then aggregating to the group level eliminates the bias if the number of individuals in each group is sufficiently large. When all of the parameter variation is within groups, $\lambda_g = 0$ and the average of λ_i converges to zero and fixed effects is consistent when $T \to \infty$.

4 Monte Carlo Simulations

4.1 Simulation Design

We simulate a model with the multilevel structure defined in (1), (2), and (3). Parameter values in the simulations are given in table 1. We set the true value of the coefficient on the lagged dependent variable equal to -0.5 because this is roughly what we expect the coefficient to equal in our application to corn acreage response to price. We randomly draw the coefficients from a normal distribution, where the standard deviations vary across our simulations. We replace values of the lagged dependent variable with -1 if the random draw is less than -1, and replaced the value with 1 if the random draw is greater than 1.

The values of y_{it} and x_{it} are set to their mean, 0, initially. The first ten simulated periods are discarded, giving eleven remaining periods. We use eleven periods since this is the length of the panel considered in our application. Discarding the first ten periods creates a realistic scenario where the initial values in the estimation are correlated with the individual-level parameters. The aggregate data are created by simply taking the average of y_{it} and x_{it} within each group.

In all of the simulations, we assume that the coefficient on the lagged dependent variable is not correlated with the coefficient on the independent variable. Therefore, after aggregating each group, the coefficient on the lagged dependent variable should approximately equal the mean coefficient within the group (Lewbel, 1994). If estimators with aggregate data deviate

from the mean coefficient, it represents bias from pooling coefficients of a heterogeneous dynamic panel rather than bias from the aggregation itself.

We consider three different estimators of the mean coefficient on the lagged dependent variable (γ) and the mean coefficient on the independent variable (β) with individual and aggregate data. We consider the grouped coefficients estimator (GC) with individual data. We consider pooled fixed effects (FE) and the Arellano-Bond estimator (A-B) with individual and aggregate data. The grouped coefficients estimator is implemented by estimating a separate regression for each group, then averaging the coefficients across groups. The Arellano-Bond estimator is the two-step GMM estimator that uses all of the available lags as instruments for the differenced equation.

4.2 Monte Carlo Results

Average estimates of γ from the Monte Carlo simulations are presented in figure 1. Results for the estimate of β are in figure 2. Each plot gives results for a different set of parameters. The first column of plots gives results for $\sigma_\gamma = 0.1$, the second column of plots gives results for $\sigma_\gamma = 0.25$, and the third column of plots gives results for $\sigma_\gamma = 0.5$. The first row of plots gives results for $\sigma_\beta = 0.2$, the second row of plots gives results for $\sigma_\beta = 0.5$, and the third row of plots gives results for $\sigma_\beta = 1$. The x-axis indicates results for different intraclass correlation coefficients (i.e., the proportion of the parameter variance that is within groups). When $\delta = 0$, none of the parameter variation is within groups. When $\delta = 1$, all of the parameter variation is within groups. Different lines on the graph indicate estimates from the alternative estimators.

Results in figure 1 show estimates of the coefficient on the lagged dependent variable. For our simulations, the grouped coefficients estimator typically has less bias than the other estimators when less than 1/3 of the parameter variation is within groups. The grouped coefficients estimator can lead to a substantial reduction in the bias over other estimators with individual data when the heterogeneity of the coefficients is large.

Fixed effects and Arellano-Bond estimates of the coefficient on the lagged dependent variable decrease when the variance of γ_i is larger, and increase when the variance of β_i is larger. The sign of the bias of fixed effects and Arellano-Bond estimates often depends on whether individual or aggregate data are used. Although fixed effects and Arellano-Bond estimates typically have a negative bias with individual data, they can have a positive bias with aggregate data when most of the parameter heterogeneity is within groups. The positive bias with aggregate data occurs because the bias from the heterogeneity of β_i is positive and increases when data are aggregated over relatively homogenous groups (see section 3.1). When all of the parameter variation is within groups ($\delta = 1$), then the aggregation reduces the parameter heterogeneity and the bias of the estimators with aggregate data is reduced, but only converges to the true parameter value when the variation in the coefficient on the lagged dependent variable is small.

Results in figure 2 show estimates of the average coefficient on the independent variable. Again, the grouped coefficients estimator gives substantial bias reduction when less than 1/3 of the parameter heterogeneity is within groups. With individual data, pooled fixed effects and Arellano-Bond estimates are biased upward. With aggregate data, the direction of the bias of fixed effects depends on the parameter values, but Arellano-Bond estimates are biased upward.

Figures 3 and 4 show the root mean squared error for the alternative estimators of γ and β. When the heterogeneity ofγ_i is small ($\sigma_\gamma = 0.1$) and less than 1/3 of the parameter variation is within groups, the performance of the grouped coefficients estimator is comparable to the Arellano-Bond estimator with individual data. When the heterogeneity of γ_i is larger ($\sigma_\gamma = 0.25$ or $\sigma_\gamma = 0.5$), then the grouped coefficients estimator has a smaller root mean squared error than the Arellano-Bond estimator when less than 2/3 of the parameter variation is within groups.

5 Application to Corn Acreage Response to Price

We illustrate the use of the grouped coefficients estimator by estimating corn acreage response to price and compare panel data estimates with field-level and county-level data in Iowa, Illinois, and Indiana. The seminal work of Nerlove (1958) interpreted the coefficient on a lagged dependent variable in agricultural supply response as capturing a partial adjustment process and adaptive expectations. A large literature applied his methods to agricultural supply response and found the long-run response was larger than the short-run response (Askari and Cummings, 1977). Eckstein (1984) showed that the same linear specification with a lagged dependent variable could be derived from a rational expectations model where the share of land devoted to one crop in the previous year affects the yield of the other crop. Eckstein's model suggests that the long-run response is smaller than the short-run response if monoculture decreases yields and increases costs.

A feature of many cropping systems is that crops are often rotated. For example, it is common in the Corn Belt to alternate between corn and soybeans from year to year on the same field. Crop rotations decrease the optimal input use (e.g. soybeans fix nitrogen that is used by the subsequent corn crop) and increase the yield in a way that cannot be substituted with market inputs. Hendricks (2011) has shown that the dynamic behavior implied by aggregating these crop rotations across heterogeneous fields implies that the long-run response to price is smaller than the short-run response.

We approximate the farmer's decision of whether to plant corn or soybeans on field i in year t with a linear probability model,

$$c_{it} = \gamma_i c_{i,t-1} + \boldsymbol{\beta}_i \mathbf{p}_{it} + \kappa_i t + \alpha_i + \varepsilon_{it}, \tag{15}$$

where c_{it} is a binary variable that equals 1 if corn is planted and 0 if soybeans are planted, $\mathbf{p}_{it}$ is a vector containing the expected revenue of corn and the expected revenue of soybeans, and $\kappa_i t$ represents incidental trends.

We assume that expected revenue is contemporaneously exogenous, such that $E\left(\varepsilon_{it}|\mathbf{p}_{it}\right) = 0$. As described in section 5.1, we construct expected revenue as the sum of a futures price, an expected basis, and an expected government payment. Certainly, the price obtained at harvest is endogenous. But the price prior to planting of a harvest-time futures contract is not endogenous as long as unobserved shocks that affect planting decisions are unknown prior to planting. Only factors that are expected prior to planting must be included in the regression for consistency. In other words, the use of a pre-planting futures price requires the standard assumption of no omitted variables for consistency. An advantage of using the pre-planting futures price is that it incorporates all of the relevant demand shocks to identify the supply response. Thus, it is more efficient than instrumental variables estimation that only exploits a subset of demand shocks.

Roberts and Schlenker (2010) argue that the futures price is endogenous and obtain larger supply elasticity estimates after instrumenting. The main reason for the difference in our conclusions is that Roberts and Schlenker (2010) estimate a supply elasticity for world production while we estimate a supply elasticity for an area within the Corn Belt. They use a futures price one year prior to delivery of harvest in the United States. While their futures price reflects expectations prior to planting in the Northern hemisphere, it also represents expectations in the middle of the growing season for the Southern hemisphere. Their futures price, therefore, is endogenous for production in the Southern hemisphere.

Let C_t equal the aggregate acres of corn and a_i equal the acres of field i. The short-run change in aggregate corn acres with respect to the expected revenue of corn is

$$\left.\frac{\partial C_t}{\partial p_{it}^c}\right|_{short-run} = \sum_i a_i \beta_i^c, \tag{16}$$

and the long-run response with respect to the expected revenue of corn is

$$\left.\frac{\partial C_t}{\partial p_{it}^c}\right|_{long-run} = \sum_i a_i \frac{\beta_i^c}{1-\gamma_i}, \tag{17}$$

where β_i^c is the coefficient on the expected revenue of corn. Equations (16) and (17) define our parameters of interest. We expect the coefficient on the lagged dependent variable, γ_i, to be negative due to crop rotation incentives; corn is less likely to be planted if soybeans were planted the previous year. We also expect γ_i to vary spatially, since monoculture is more prominent in certain locations. The short-run response to price, β_i, may also vary spatially as relative yields of corn and soybeans vary across soil types.

5.1 Field-Level Data Description

Our field-level crop data are classified from satellite imagery and released as the Cropland Data Layer by the National Agricultural Statistics Service (NASS).[6] The Cropland Data Layer is an image of an entire state with a crop or land use classification code corresponding to each pixel, where pixels are less than one acre in size (roughly 64 meters x 64 meters). The Cropland Data Layer is available in Illinois for the period 1999–2010 and Iowa and Indiana for 2000–2010. NASS conducts an accuracy assessment of the Cropland Data Layer for each state. On average, the probability that the Cropland Data Layer correctly classifies corn or soybeans is roughly 95%.

We use Common Land Unit (CLU) boundaries from the Farm Service Agency to approximate "field" boundaries and choose a point near the centroid of the common land unit as our unit of analysis.[7] Our panel of crop data corresponds to the land-use classification of the pixels at each of these points over time. Our dependent variable with field-level data is a binary variable that indicates corn or soybeans. The Farm Service Agency states that the common land unit data layer contains digitized field boundaries, where fields are defined as

[6]The Cropland Data Layer can be viewed and downloaded at `http://nassgeodata.gmu.edu/CropScape/`.

[7]We do not want to use each pixel as a separate unit of analysis for the econometric model for at least two reasons: i) pixels are more likely to be misclassified at field boundaries (De Wit and Clevers, 2004), and ii) a farmer's crop decision is at the field level, not the pixel level. The center of a CLU is also problematic because CLUs are often split into two fields, so the center point may not be classified from the same field each year. To avoid these concerns, we use points that are diagonally offset from the centroid of each CLU. The distance that the point is offset from the centroid is in proportion to the size of the CLU and the direction is chosen randomly.

"agricultural land that is delineated by natural and man-made boundaries such as road ways, tree lines, waterways, fence lines, etc." (Farm Service Agency, 2007). Many of the common land units that are extremely small are likely to be gullies, waterways, or farmsteads so only points corresponding to common land units larger than 15 acres are included in the empirical analysis.

We only keep observations (field-year pairs) that were classified as corn or soybeans for the econometric analysis. Double-cropped wheat and soybeans are included in the soybeans category. To keep our presentation of results concise, we do not model the conversion of land between corn or soybeans and other land uses since the acreage response at this margin is a small proportion of the total acreage response in this region (Hendricks, 2011).

In our econometric model, the right-hand side regressor is expected revenue rather than simply expected price because the government distorts the revenue that farmers receive when corn and soybean prices are low. Expected crop revenues per bushel are the sum of a futures price, an expected basis (i.e., the difference between the cash and futures prices), and an expected loan deficiency payment.

For corn, the futures price is the average price in January–March (prior to planting) of a December futures contract (after harvest). For soybeans, the futures price is the average price in January–March of a November futures contract. Futures price data are obtained from the Commodity Research Bureau. The basis in March is used as the expectation for the basis at harvest. Spot prices in March were purchased for 93 market locations in the three states for corn and 90 locations for soybeans from Cash Grain Bids, Inc. The expected basis is then interpolated to every point in the crop dataset using inverse distance weighting. Loan deficiency payments from U.S. government are triggered when market prices fall below the loan rate. We assume the loan rate truncates the distribution of anticipated prices and construct the expected loan deficiency payment by calculating the difference in the truncated and untruncated expected market price—Chavas and Holt (1990) and Barr et al. (2011), among others, follow a similar approach. We assume anticipated prices follow a lognormal

distribution, where the untruncated mean is the futures price and untruncated standard deviation is the average implied volatility in January–March of the December options for corn and November options for soybeans obtained from the Commodity Research Bureau.

Soil data are obtained from the Soil Survey Geographic (SSURGO) database released by the Natural Resources Conservation Service (NRCS). In the econometric model, we estimate different coefficients for different soil regimes. We define soil regimes using the soil taxonomy from NRCS. The soil taxonomy does not impose a particular judgment about the production capability of a soil, but merely defines the fields that have soils with similar characteristics.

The soil taxonomy has six classification hierarchies. For fields that planted corn or soybeans in the three states, there are 1,186 unique taxonomy classifications. We reduce the dimension of the taxonomy to 59 different soil regimes so that there are at least 1,000 observations per year in each soil regime.

5.2 County-Level Data Description

Our county-level crop data are from the official county-level crop acreage data reported by the National Agricultural Statistics Service. NASS uses the Cropland Data Layer to inform its county-level estimates, but also uses other survey data. County-level planted acreage for corn and soybeans in Iowa, Illinois, and Indiana are available from the 1970s and harvested acreage are available from the 1920s. For the purposes of this paper, we restrict our analysis of county-level data to the same period that data from the Cropland Data Layer were available. We only use counties that had acreage data available for every year for the period 1999-2010 for Illinois and 2000-2010 for Iowa and Indiana.[8]

For the county-level data, our dependent variable is the acres planted to corn divided by the average number of acres planted to corn or soybeans for that county during the sample period. We need to normalize the acres planted to corn for each county since the acreage

[8]There were 11 counties that only had data available for a portion of the sample period. If we include these counties in the analysis below, then the fixed effects estimate of the coefficient on the lagged dependent variable is biased upward even further than the estimate we report.

depends largely on the size of the county. We divide by the average number of acres planted to corn or soybeans during the period so that the denominator does not change from year to year due to changes in prices. We also considered estimating the model in logs, but chose to specify the dependent variable as a share so that it most closely reflected aggregation of the linear probability model. Estimating the model in logs does not substantially affect the results. We obtain county-level expected revenues by calculating the average expected revenues within each county from the field-level data.

Descriptive statistics for the field-level and county-level data are given in table 2. The samples include Illinois crop data for the period 1999-2010 and Iowa and Indiana crop data for the period 2000-2010. A nonmissing "observation" is a field (or county) in a particular year that has a nonmissing dependent variable in the current year and the previous year. We lose one year of data since we include a lagged dependent variable. For the field-level data, observations may also be missing in other years due to cloud cover or because the field was classified as a land use other than corn or soybeans. There are 8.4 million observations in the field-level dataset and 2,917 observations in the county-level dataset.

5.3 Results

We present our coefficient estimates with alternative estimators and alternative forms of aggregation in figures 5, 6, and 7. The top panel of each figure gives estimates from field-level data using the Cropland Data Layer. The middle panel gives estimates from county-level data using official NASS county-level data. The bottom panel gives the estimates where we randomly assign each field to a group, then average the data within each group for each year. We create 100 random groups of fields. The dependent variable for the field-level data is a binary variable indicating corn or soybeans. The dependent variable for the county-level data is the acres planted to corn divided by the average number of acres planted to corn or soybeans for that county during the sample period. The aggregate data simply

represent averages of the field-level data so that coefficients estimates across the alternative aggregations are directly comparable.[9]

We consider the grouped coefficients estimator with field level data. We define groups of homogenous fields using the soil taxonomy and the observed crop transitions during the period. Fields are divided into 59 different soil regimes using the soil taxonomy as described in section 5.1. We define groups within the soil regimes of those fields that i) always planted corn after soybeans and soybeans after corn, ii) always planted corn after corn, iii) always planted soybeans after soybeans, and iv) fields that transitioned between corn and soybeans at least once but also planted corn after corn or soybeans after soybeans at least once. About 32% of acres were always planted to corn after soybeans or to soybeans after corn during the sample period, about 1% of acres were always planted to corn after corn, and about 0.5% of acres were always planted to soybeans after soybeans.[10]

We calculate two-way cluster-robust, bootstrap standard errors of the average coefficients across group-specific regressions, where we cluster by year and field and stratify the bootstrap samples by our groups. A description of two-way cluster-robust standard errors is given by Cameron, Gelbach, and Miller (2011). The bootstrap samples are drawn independently within each group (stratum) so that each group is represented in each bootstrap sample. One-way cluster covariance matrices are estimated by resampling clusters with replacement. The two-way cluster covariance matrix is the sum of the covariance matrix clustered by year and the covariance matrix clustered by field minus the covariance matrix with each field, year pair drawn independently.

We also consider pooled OLS, pooled fixed effects, and the Arellano-Bond estimator for the field and county-level data. Each of the models includes a trend variable. We do

[9]Since counties differ in size, we do not consider aggregate corn acres as a dependent variable. We did estimate the county-level model in logarithms. Elasticity estimates were generally smaller in the logarithmic model, but are not reported here.

[10]For fields that were always planted to corn after soybeans or to soybeans after corn, the OLS estimate of the coefficient on the lagged dependent variable is -1 and the estimates of the coefficients on prices are 0. For fields that were always planted to corn after corn or soybeans after soybeans, the coefficient on the lagged dependent variable is not identified but the coefficients on prices are 0.

not control for year fixed effects because our variation in price occurs primarily between years, and thus we would lose the variation of greatest importance for this application. Our Arellano-Bond estimator uses two-step GMM with every lag available as an instrument for the difference of the lagged dependent variable. The grouped coefficients estimator is not appropriate for the county-level data since counties cannot be grouped into large, homogenous groups.[11]

Standard errors for pooled OLS, pooled fixed effects, and Arellano-Bond are clustered at the crop-reporting district level. There are 27 districts in the three states. Bester, Conley, and Hansen (2011) propose clustering the standard errors in large spatial blocks in order to account for spatial dependence in the errors. However, this clustering approach is not appropriate for the grouped coefficients estimator.

We estimate the model with randomly assigned group aggregation using pooled OLS. The aggregation is intended to remove parameter heterogeneity—including heterogeneity of the intercept—such that fixed effects are not necessary for the dynamic model.

Figure 5 shows the estimates of the coefficient on the lagged dependent variable across specifications. The difference between estimates using field and county-level data is remarkable. With field-level data we obtain a coefficient on the lagged dependent variable less than -0.54 with every estimator. The grouped coefficients estimate of the average coefficient on the lagged dependent variable is -0.59. With county-level data we obtain a positive coefficient on the lagged dependent variable with pooled OLS and pooled fixed effects. The Arellano-Bond estimator with county-level data gives a negative coefficient on the lagged dependent variable, but is still substantially different from the estimates with field level data.

Some people may have a vague notion that the aggregate data capture a partial adjustment process while the field-level data capture the dynamics from crop rotations. In other words, it may be that the estimates with aggregate data are capturing a different dynamic process that is of interest for estimating the long-run response to price. However, this notion

[11]Results using the grouped coefficients estimator where counties were grouped by crop-reporting districts gave results similar to pooled OLS with county-level data.

is rejected by the fact that aggregating the data into random groups and estimating with pooled OLS gives an estimate on the lagged dependent variable that is similar to estimates with field-level data. Aggregation itself does not substantially affect the dynamics, but rather the form of aggregation can substantially affect the bias of standard panel estimators (i.e., aggregating by groups with substantial between-group heterogeneity).

Figure 6 shows the estimates of the coefficient on expected revenue per bushel of corn (i.e., expected market price plus expected government payments), and figure 7 shows the estimates of the coefficient on expected revenue per bushel of soybeans across specifications. The grouped coefficients estimate is similar to pooled fixed effects and Arellano-Bond with field-level data. Pooled OLS at the field level underestimates the effect of prices on planting decision. Results across alternative estimators are more variable with county-level data than with field-level data. With county-level data, pooled OLS gives the largest coefficient on revenue and the Arellano-Bond estimator gives the smallest coefficient on revenue.

Although the coefficient on the lagged dependent variable was similar with field-level data and randomly aggregated data, the coefficients on revenue are much small with the randomly aggregated data. The coefficients on revenue with the randomly aggregated data are similar to the coefficients on revenue using pooled OLS with field-level data.

Figure 8 show the long-run elasticity of corn acreage with respect to the price of corn and figure 9 shows the long-run elasticity of corn acreage with respect to the price of soybeans. Note that we have converted the marginal effects into elasticities for these two figures. The long-run marginal effect for the grouped coefficients estimator is calculated as a weighted average of the group-specific long-run marginal effects, where weights correspond to the acres in each group, then converted into an elasticity.

Using the grouped coefficients estimator, we obtain a long-run own-price elasticity of 0.23 and long-run cross-price elasticity of -0.17. We find that the coefficient on the lagged dependent variable is positively correlated with the coefficients on the price of corn and negatively correlated with the coefficient on the price of soybeans. For example, the fields

that were always in a corn-soybean rotation have a coefficient on the lagged dependent variable of -1 and the coefficients on prices are 0, while other fields have a larger coefficient on the lagged dependent variable (i.e., $\gamma_i > -1$) and a larger coefficient on the price of corn (i.e., $\beta_i^c > 0$). If we ignore this correlation and calculate the long-run elasticity using the average coefficient on the lagged dependent variable and the average coefficients on prices, then we obtain a long-run own-price elasticity of 0.20 and a long-run cross-price elasticity of -0.15.

The long-run own-price elasticity with the grouped coefficients estimator is similar to the fixed effects and Arellano-Bond estimates, but the cross-price elasticity is larger (in absolute value). Estimates of the long-run elasticities with pooled OLS and pooled fixed effects with county-level data differ substantially from the estimates with field-level data. We do not display the pooled OLS estimate of the long-run elasticities in figures 8 and 9 since they are so large that they would make the figures difficult to read (own-price elasticity of 3.3 and cross-price elasticity of -4.6). The pooled fixed effects elasticity estimate with county-level data are roughly twice as large as the grouped coefficients estimate because of the difference in the sign of the coefficient on the lagged dependent variable.

The Arellano-Bond long-run elasticities with county-level data are similar to the grouped coefficients elasticity estimates. The two estimators, however, imply different dynamics of the supply response. The Arellano-Bond estimate with county-level data indicates a smaller response in the short-run and a smaller difference between the short-run and long-run response than the grouped coefficients estimator.

6 Conclusion

Our paper illustrates the importance of considering the multilevel structure of dynamic panels. We propose the use of a grouped coefficients estimator to reduce the bias in dynamic panels with coefficient heterogeneity, common factors, and small T. The grouped coefficients

estimator reduces the bias of mean coefficient estimates if groups of relatively homogenous individuals can be defined. In Monte Carlo simulations, we show that the grouped coefficients estimator can reduce the bias compared to fixed effects and Arellano-Bond estimates.

We also describe how the bias of fixed effects differs with individual and aggregate data when there are heterogeneous coefficients. Monte Carlo simulations show that there can be large differences in the bias of fixed effects and Arellano-Bond estimates with individual and aggregate data depending on whether the data are aggregated over relatively homogenous individuals or relatively heterogeneous individuals.

In an application to estimating corn acreage response to price, we illustrate the use of the grouped coefficients estimator by defining groups of fields with similar soils and observed crop transition histories. Estimates of the coefficient on the lagged dependent variable are dramatically different depending on whether field or county-level data are used. Field-level data indicate that the long-run response to price is smaller than the short-run response to price—consistent with the dynamics of supply that are implied by crop rotations. Fixed effects estimates with county-level data indicate that the long-run response to price is larger than the short-run response to price which has often been interpreted as a partial adjustment process.

References

Anderson, T.W., and C. Hsiao. 1982. "Formulation and Estimation of Dynamic Models using Panel Data." *Journal of Econometrics* 18:47–82.

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

Arellano, M., and O. Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-Components Models." *Journal of Econometrics* 68:29–51.

Askari, H., and J.T. Cummings. 1977. "Estimating Agricultural Supply Response with the Nerlove Model: A Survey." *International Economic Review* 18:257–292.

Barr, K.J., B.A. Babcock, M.A. Carriquiry, A.M. Nassar, and L. Harfuch. 2011. "Agricultural Land Elasticities in the United States and Brazil." *Applied Economic Perspectives and Policy* 33:449–462.

Bester, C.A., T.G. Conley, and C.B. Hansen. 2011. "Inference with Dependent Data using Cluster Covariance Estimators." *Journal of Econometrics* In Press, Accepted Manuscript.

Bester, C.A., and C.B. Hansen. 2009. "Grouped Effects Estimators in Fixed Effects Models." Working Paper.

Cameron, A.C., J.B. Gelbach, and D.L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business and Economic Statistics* 29:238–249.

Chavas, J.P., and M.T. Holt. 1990. "Acreage Decisions under Risk: The Case of Corn and Soybeans." *American Journal of Agricultural Economics* 72:529–538.

De Wit, A.J.W., and J.G.P.W. Clevers. 2004. "Efficiency and Accuracy of Per-Field Classification for Operational Crop Mapping." *International Journal of Remote Sensing* 25:4091–4112.

Deaton, A. 1988. "Quality, Quantity, and Spatial Variation of Price." *The American Economic Review* 78:418–430.

Eckstein, Z. 1984. "A Rational Expectations Model of Agricultural Supply." *The Journal of Political Economy* 92:1–19.

Farm Service Agency. 2007. "Common Land Unit Metadata." Available with original data.

Granger, C.W.J. 1987. "Implications of Aggregation with Common Factors." *Econometric Theory* 3:208–222.

—. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of Econometrics* 14:227–238.

Hahn, J., and G. Kuersteiner. 2002. "Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects when Both n and T are Large." *Econometrica* 70:1639–1657.

Hendricks, N.P. 2011. "The Dynamics and Spatial Heterogeneity of Crop Acreage Response to Price: Problems of Aggregation and Pooling." PhD dissertation, University of California, Davis.

Hsiao, C., M.H. Pesaran, and A.K. Tahmiscioglu. 1999. "Bayes Estimation of Short-Run Coefficients in Dynamic Panel Data Models." In C. Hsiao, K. Lahiri, L.-F. Lee, and M. H. Pesaran, eds. *Analysis of Panels and Limited Dependent Variable Models*. Cambridge, UK: Cambridge University Press.

Kiviet, J.F. 1995. "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models." *Journal of Econometrics* 68:53–78.

Lancaster, T. 2000. "The Incidental Parameter Problem since 1948." *Journal of Econometrics* 95:391–413.

Lewbel, A. 1994. "Aggregation and Simple Dynamics." *The American Economic Review* 84:905–918.

Nauges, C., and A. Thomas. 2003. "Consistent Estimation of Dynamic Panel Data Models with Time-Varying Individual Effects." *Annals of Economics and Statistics*, pp. 53–75.

Nerlove, M. 1958. *The Dynamics of Supply: Estimation of Farmers' Response to Price*. Baltimore, MD: The Johns Hopkins Press.

Nickell, S. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49:1417–1426.

Pesaran, M.H., and R. Smith. 1995. "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels." *Journal of Econometrics* 68:79–113.

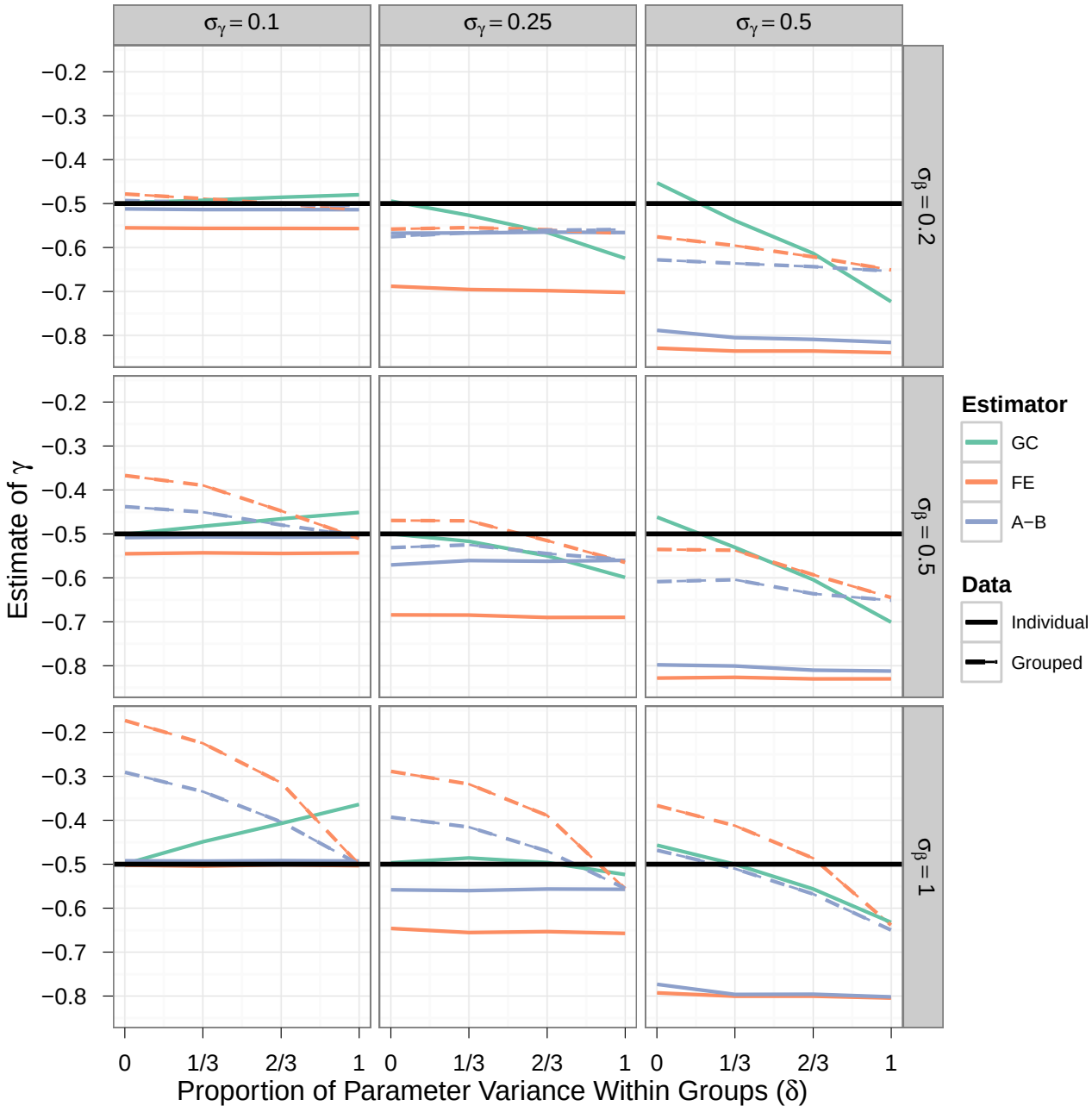
Phillips, P.C., and D. Sul. 2007. "Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence." *Journal of Econometrics* 137:162–188.

Roberts, M.J., and W. Schlenker. 2010. "Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate." Working Paper No. 15921, National Bureau of Economic Research.

Robertson, D., and J. Symons. 1992. "Some Strange Properties of Panel Data Estimators." *Journal of Applied Econometrics* 7:175–189.

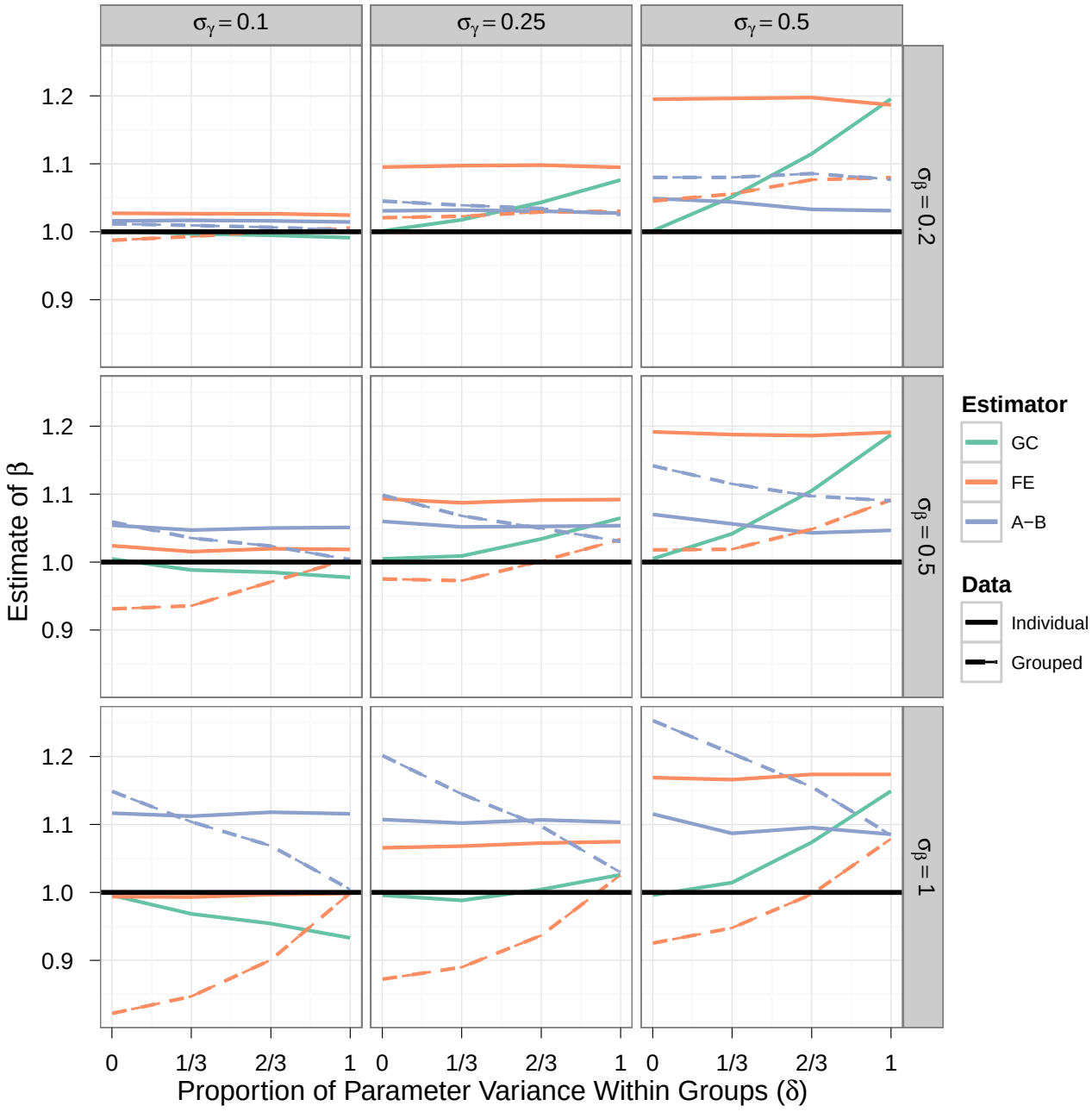
Sarafidis, V., and D. Robertson. 2009. "On the Impact of Error Cross-Sectional Dependence in Short Dynamic Panel Estimation." *Econometrics Journal* 12:62–81.

Figure 1: Monte Carlo Results: Estimates of γ



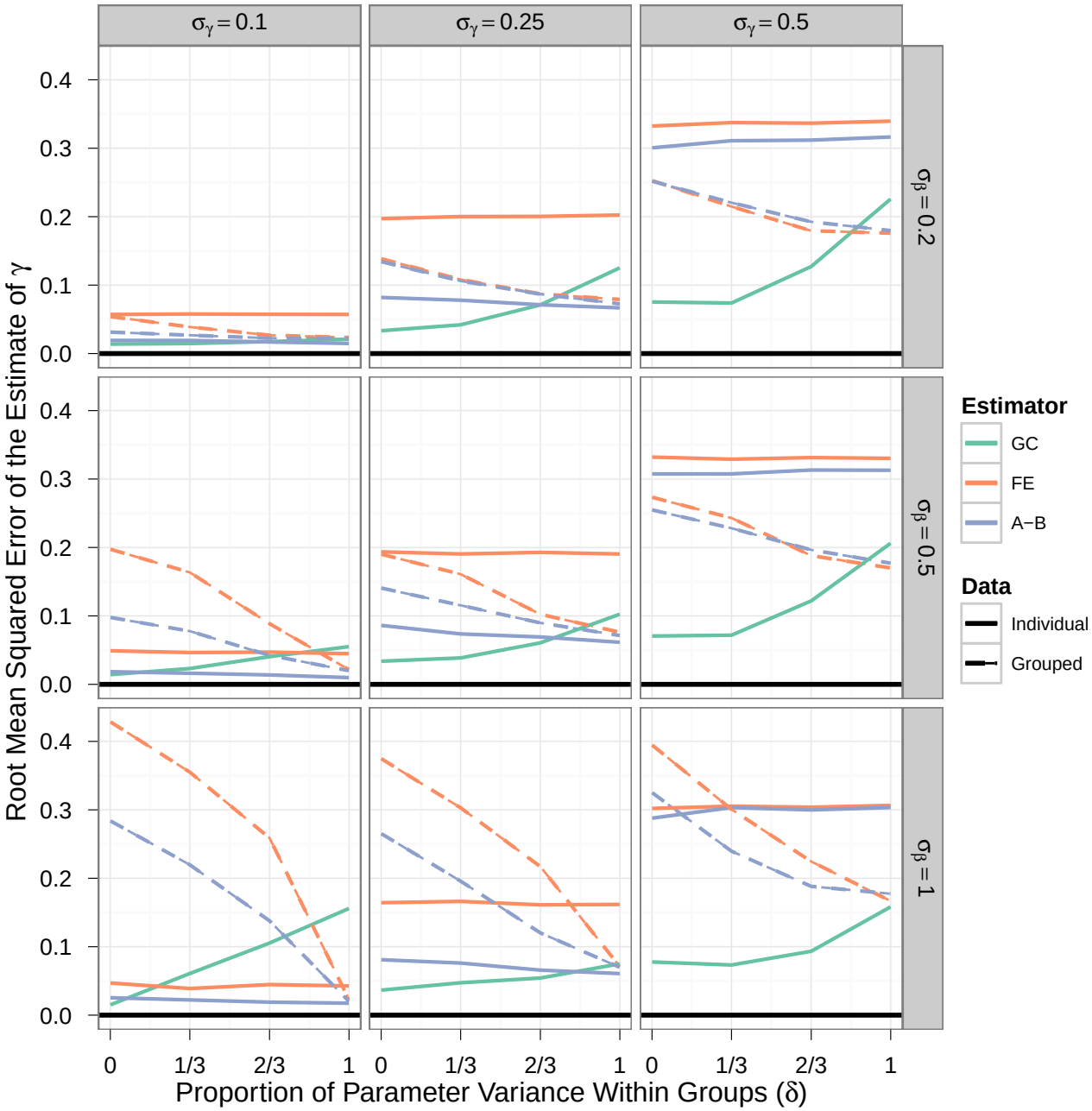
Notes: GC=Grouped Coefficients estimator, FE=Pooled Fixed Effects estimator, and A-B=Arellano-Bond estimator. The figure presents Monte Carlo simulations for the estimate of the mean coefficient on the lagged dependent variable for the equation $y_{it} = \gamma_i y_{i,t-1} + \beta_i x_{it} + \alpha_i + \varepsilon_{it}$, where γ_i, β_i, and α_i have a multilevel structure.

Figure 2: Monte Carlo Results: Estimates of β



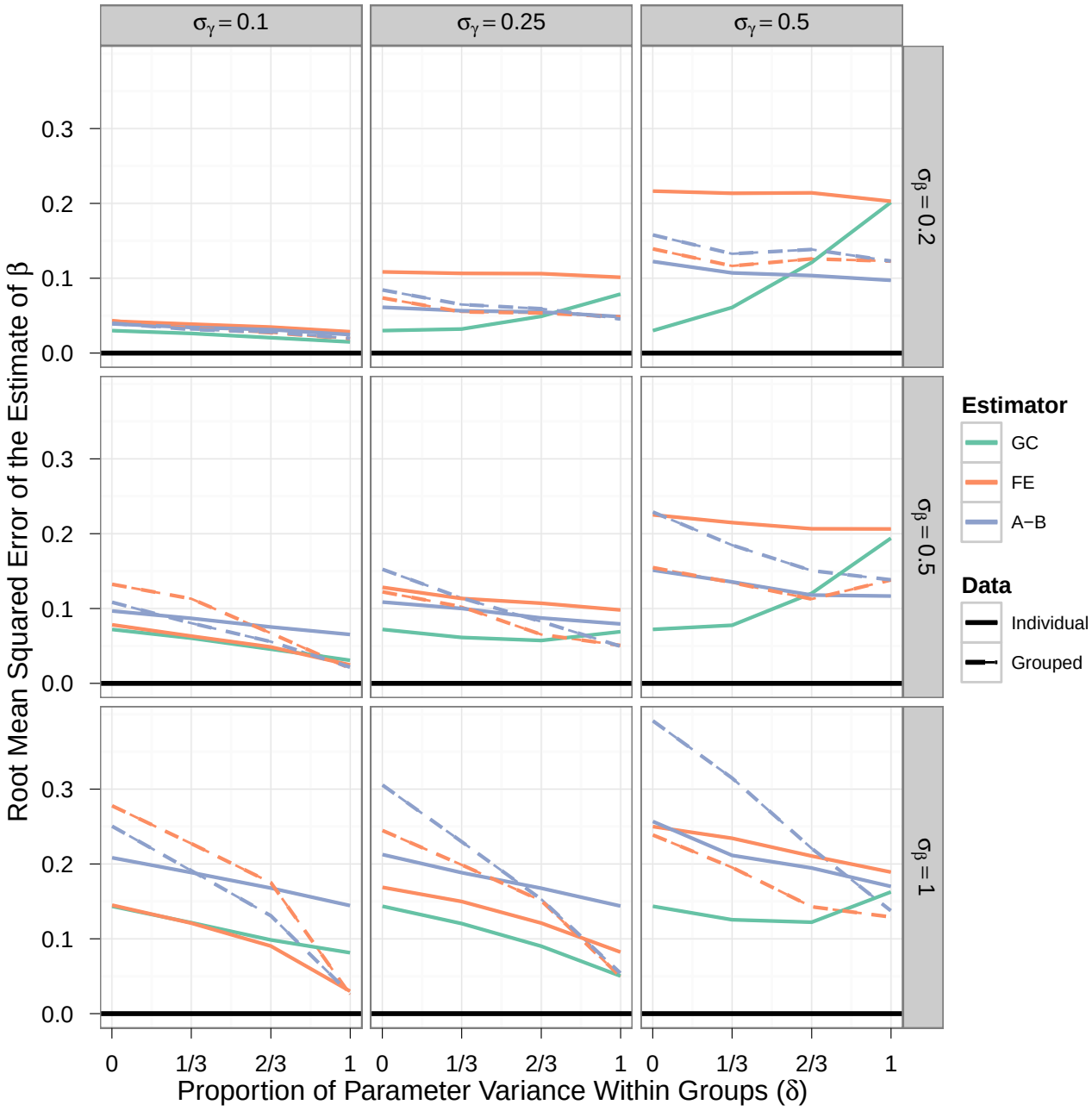
Notes: GC=Grouped Coefficients estimator, FE=Pooled Fixed Effects estimator, and A-B=Arellano-Bond estimator. The figure presents Monte Carlo simulations for the estimate of the mean coefficient on the independent variable for the equation $y_{it} = \gamma_i y_{i,t-1} + \beta_i x_{it} + \alpha_i + \varepsilon_{it}$, where γ_i, β_i, and α_i have a multilevel structure.

Figure 3: Monte Carlo Results: Root Mean Squared Error of Estimates of γ



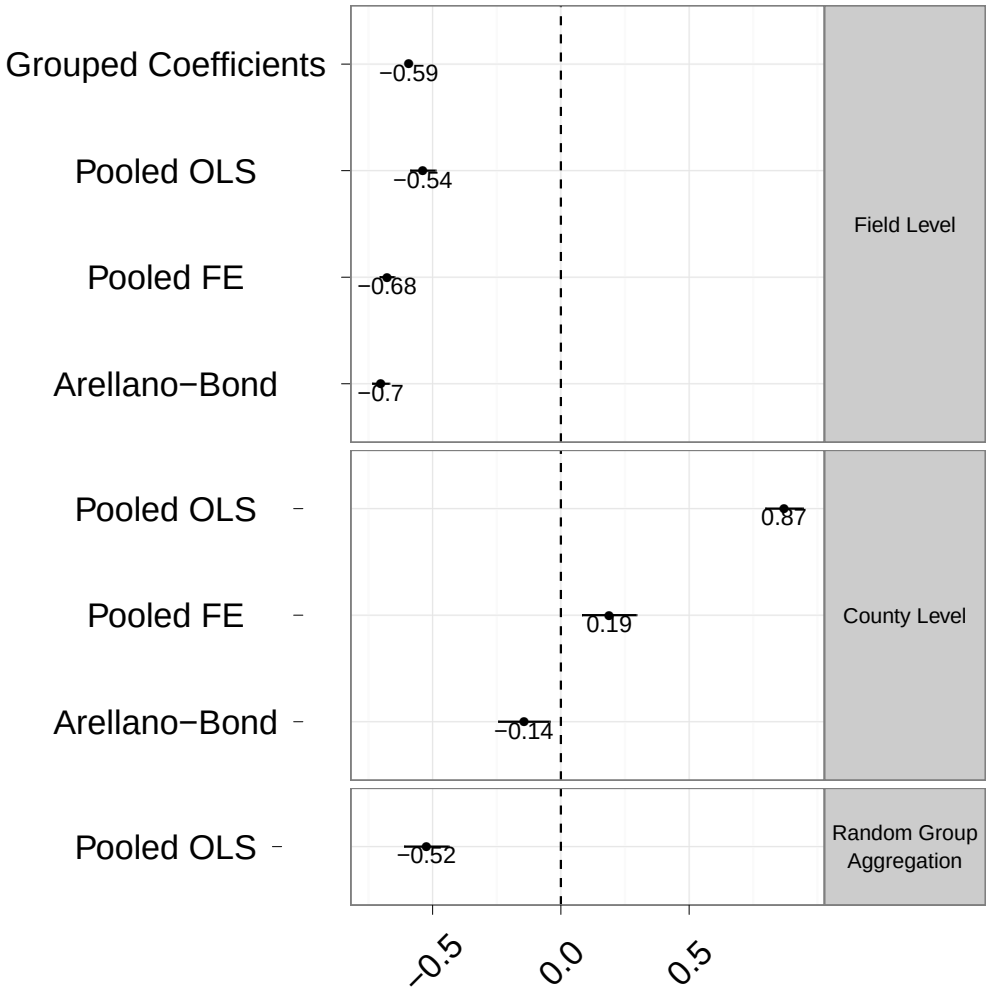
Notes: GC=Grouped Coefficients estimator, FE=Pooled Fixed Effects estimator, and A-B=Arellano-Bond estimator. The figure presents Monte Carlo simulations for the root mean squared error of the the estimate of the mean coefficient on the lagged dependent variable for the equation $y_{it} = \gamma_i y_{i,t-1} + \beta_i x_{it} + \alpha_i + \varepsilon_{it}$, where γ_i, β_i, and α_i have a multilevel structure.

Figure 4: Monte Carlo Results: Root Mean Squared Error of Estimates of β



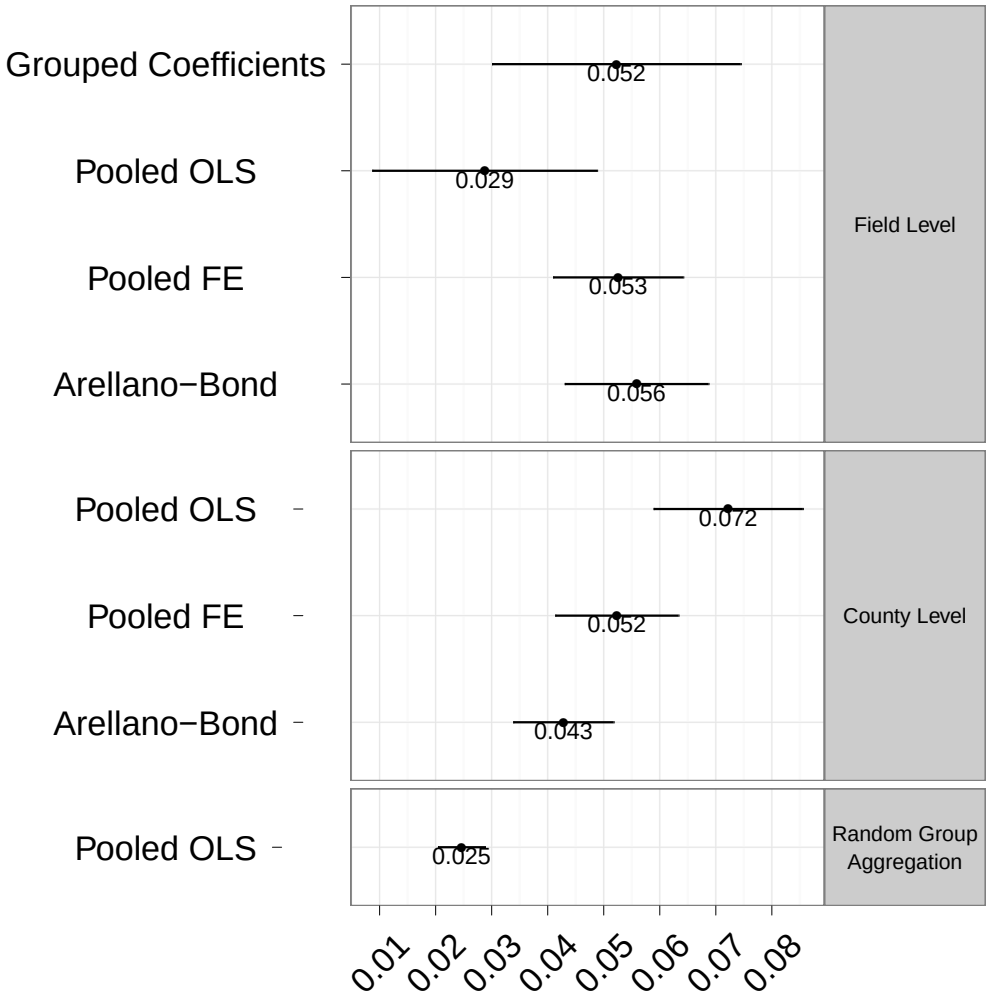
Notes: GC=Grouped Coefficients estimator, FE=Pooled Fixed Effects estimator, and A-B=Arellano-Bond estimator. The figure presents Monte Carlo simulations for the root mean squared error of the estimate of the mean coefficient on the independent variable for the equation $y_{it} = \gamma_i y_{i,t-1} + \beta_i x_{it} + \alpha_i + \varepsilon_{it}$, where γ_i, β_i, and α_i have a multilevel structure.

Figure 5: Estimates of Coefficient on Lagged Dependent Variable for Corn Planting Response to Price with Alternative Estimators and Alternative Aggregation



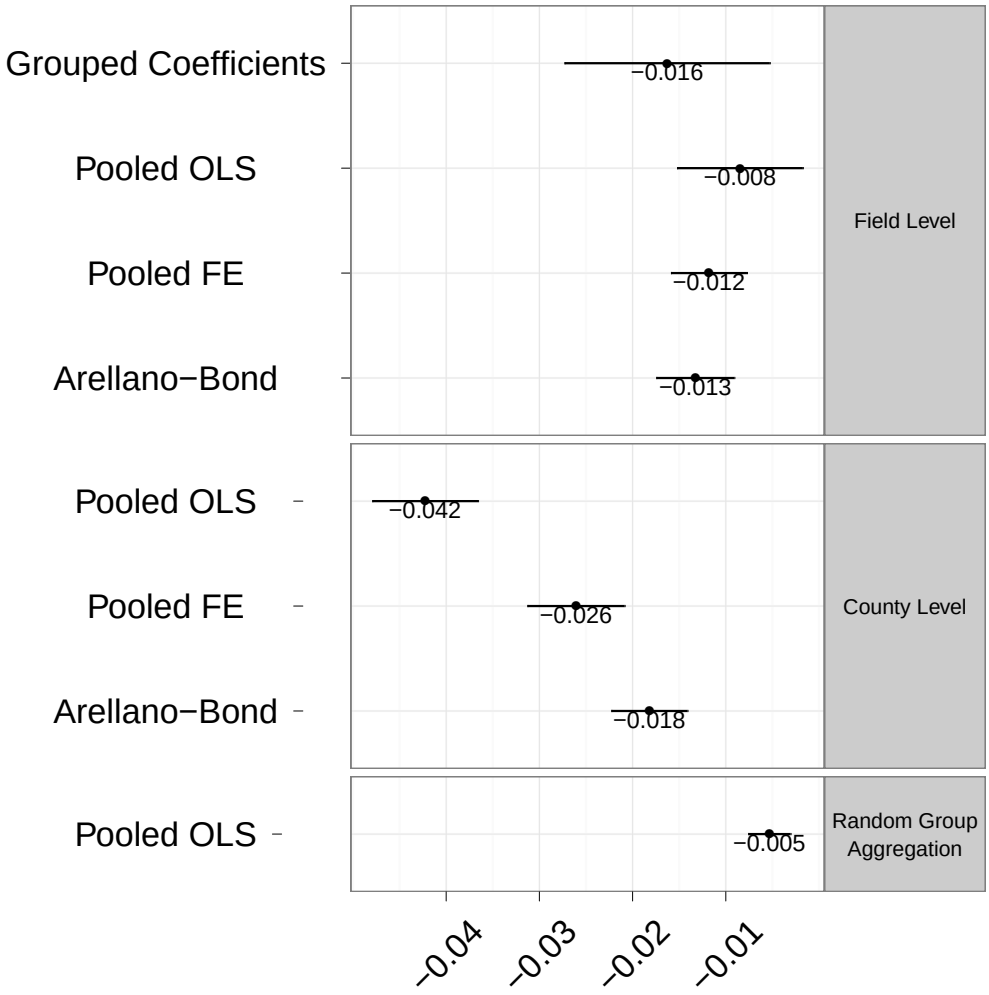
Notes: The panels in the figure give estimates from field-level data, county-level data, and randomly aggregated data. The dependent variable for the field-level data is a binary variable indicating corn or soybeans. The dependent variable for the county-level data is the acres planted to corn divided by the average number of acres planted to corn or soybeans for that county during the sample period. The dependent variable for the randomly aggregated data is the average of the corn binary variables within each group. The regressors include a lagged dependent variable, expected revenue per bushel of corn, expected revenue per bushel of soybeans, and a time trend. Standard errors for the grouped coefficients estimator are two-way (field and year) cluster-robust, bootstrap standard errors. Standard errors for pooled OLS, fixed effects, and Arellano-Bond with field and county-level data are clustered at the crop-reporting district level. Standard errors for OLS with randomly aggregated data are not heteroskedastic consistent.

Figure 6: Estimates of Coefficient on Corn Revenue per Bushel for Corn Planting Response to Price with Alternative Estimators and Alternative Aggregation



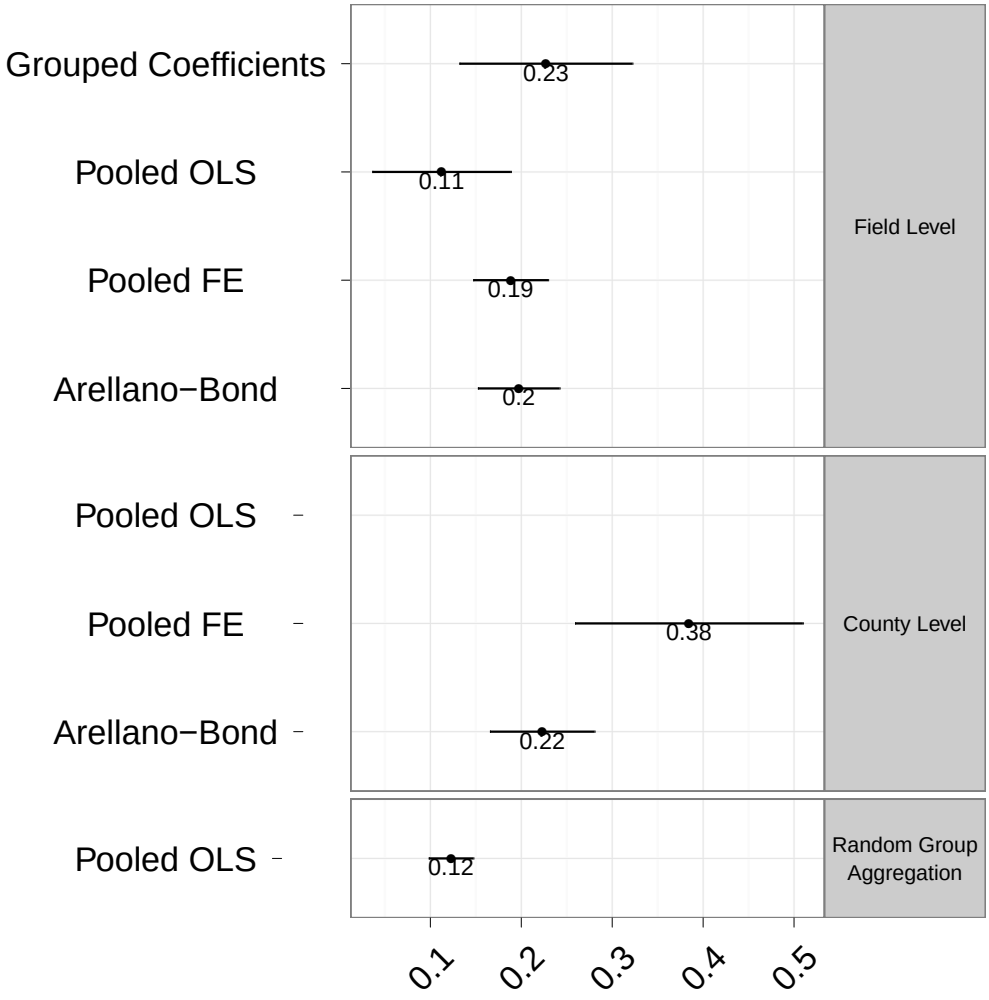
Notes: The panels in the figure give estimates from field-level data, county-level data, and randomly aggregated data. The dependent variable for the field-level data is a binary variable indicating corn or soybeans. The dependent variable for the county-level data is the acres planted to corn divided by the average number of acres planted to corn or soybeans for that county during the sample period. The dependent variable for the randomly aggregated data is the average of the corn binary variables within each group. The regressors include a lagged dependent variable, expected revenue per bushel of corn, expected revenue per bushel of soybeans, and a time trend. Standard errors for the grouped coefficients estimator are two-way (field and year) cluster-robust, bootstrap standard errors. Standard errors for pooled OLS, fixed effects, and Arellano-Bond with field and county-level data are clustered at the crop-reporting district level. Standard errors for OLS with randomly aggregated data are not heteroskedastic consistent.

Figure 7: Estimates of Coefficient on Soybean Revenue per Bushel for Corn Planting Response to Price with Alternative Estimators and Alternative Aggregation



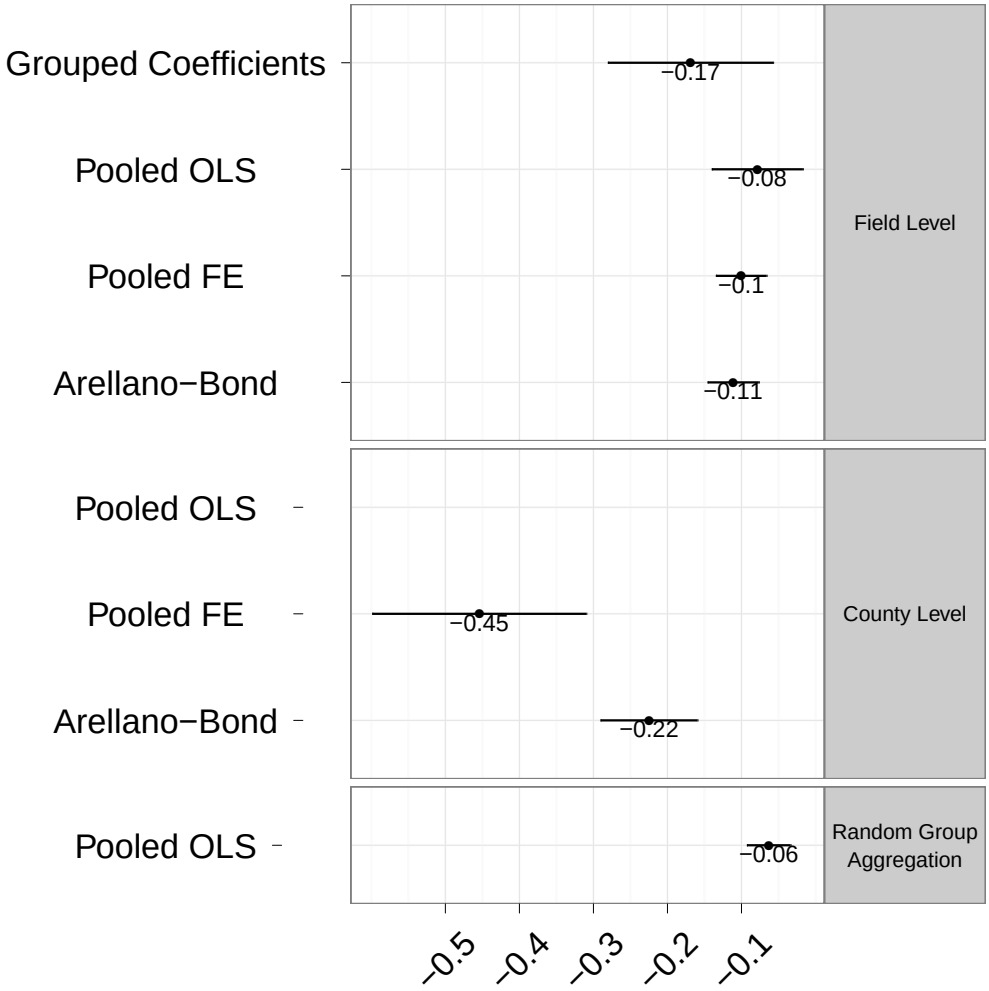
Notes: The panels in the figure give estimates from field-level data, county-level data, and randomly aggregated data. The dependent variable for the field-level data is a binary variable indicating corn or soybeans. The dependent variable for the county-level data is the acres planted to corn divided by the average number of acres planted to corn or soybeans for that county during the sample period. The dependent variable for the randomly aggregated data is the average of the corn binary variables within each group. The regressors include a lagged dependent variable, expected revenue per bushel of corn, expected revenue per bushel of soybeans, and a time trend. Standard errors for the grouped coefficients estimator are two-way (field and year) cluster-robust, bootstrap standard errors. Standard errors for pooled OLS, fixed effects, and Arellano-Bond with field and county-level data are clustered at the crop-reporting district level. Standard errors for OLS with randomly aggregated data are not heteroskedastic consistent.

Figure 8: Estimates of the Long-Run Elasticity of Corn Acreage Response to the Price of Corn with Alternative Estimators and Alternative Aggregation



Notes: The figure gives estimates of the long-run corn acreage elasticity with respect to the price of corn from field-level data, county-level data, and randomly aggregated data across alternative estimators. Estimates of the long-run elasticity using pooled OLS with county-level data are omitted because they are substantially larger than the other estimates and make the figure difficult to read.

Figure 9: Estimates of the Long-Run Elasticity of Corn Acreage Response to the Price of Soybeans with Alternative Estimators and Alternative Aggregation



Notes: The figure gives estimates of the long-run corn acreage elasticity with respect to the price of soybeans from field-level data, county-level data, and randomly aggregated data across alternative estimators. Estimates of the long-run elasticity using pooled OLS with county-level data are omitted because they are substantially larger than the other estimates and make the figure difficult to read.

Table 1: Parameters for Monte Carlo Simulations

Parameter	**Value(s)**
Parameters Constant across Simulations	
γ	-0.5
β	1.0
α	0
σ_α	0.5
σ_ε	2
μ_i	0
ρ	0.8
σ_u	0.75
ϕ	0.25
N	5,000
N_g	100
T	11
Number of Reps	500
Parameters Varying across Simulations	
σ_γ	$\{0.1, 0.25\}$
σ_β	$\{0.5, 1.0\}$
δ	$\{0, 1/3, 2/3, 1\}$

Table 2: Descriptive Statistics of Estimation Samples

	Data Sample	
Statistic	Field Level	County Level
Proportion Corn	0.53	0.53
Avg Expected Revenue of Corn ($/bu)	3.19	3.17
Avg Expected Revenue of Soybeans ($/bu)	7.56	7.52
Nonmissing Observations	8,386,395	2,917

Notes: The proportion corn is simply the average of the dependent variable in the respective dataset.