The Geography of Nonresponse
Can spatial econometric techniques improve survey weights for nonresponse?

Christoph Zangger\textsuperscript{1,2}

\textsuperscript{1}LMU Munich
\textsuperscript{2}University of Zurich

July 23, 2021
Background & Motivation

Figure 1: Nonresponse a national survey in Switzerland, 2018 (N ≈ 25’900)
Non-response is geographically clustered

Reason:

Segregation
- Different people end up in different geographical contexts
- (Budget) constraints (Clark and Lisowski, 2017; Schaake, Burgers, and Mulder, 2014)
- Housing preferences (Bruch and Mare, 2006; Ibraimovic and Hess, 2018)
Existing research

- GIS for mapping patterns of nonresponse (Hansen et al., 2007)
- Identification of geographic areas with hard to reach segments of the population (Low Response Score – Abbott and Compton, 2014; Erdman and Bates, 2017)
- Geocoded census characteristics as aggregated paradata for weighting survey response (Biemer and Peytchev, 2013; Olson, 2013)

⇒ Why not make use of the spatial dependence more directly?!
⇒ **Idea: Nearby units’ response status as a predictor for a particular unit’s response status**
⇒ Spatial Econometric models
Response probabilities

Response probabilities SAR (probit case):

\[
\phi_i(x_i) = p \left( [(\mathbb{I} - \rho W)^{-1} X \beta]_i + [(\mathbb{I} - \rho W)^{-1} \varepsilon]_i > 0 \right) \\
= \Phi \left\{ [(\mathbb{I} - \rho W)^{-1} X \beta]/\sigma_u \right\},
\]

where \( u = (\mathbb{I} - \rho W)^{-1} \varepsilon \).

SEM (probit case):

\[
\phi_i(x_i) = \Phi \left\{ X \beta/\sigma_u \right\}
\]

Inverse probability weights:

\[
\omega_i = \frac{1}{\phi_i(x_i)}
\]

⇒ e.g., weighted average: \( \bar{y} = \frac{\sum_{i=1}^{n} y_i \omega_i}{\sum_{i=1}^{n} \omega_i} \)
How to proceed?

Model estimation

Predict response probabilities with:
- \( oX1 \) and \( oX2 \)
- aggregated paradata \( u\bar{X}1 \) (e.g., share of people with tertiary degree in census districts)
- different models:
  - GLM
  - SEM
  - SAR
  - GLMM
  - SEM with paradata
  - SAR with paradata
Simulation

DGP

\[
\phi(X) \sim Bin(1, \pi), \quad \pi = \frac{\exp(X\beta)}{(1 + \exp(X\beta))}, \text{ where} \tag{4}
\]

\[
X = (1^T \quad uX_1 \quad oX_1 \quad oX_2),
\]

\[
uX_1 \sim (I - \rho W)^{-1} \times Bin(5, 0.2),
\]

\[
oX_1 \sim (I - \rho W)^{-1} \times Bin(1, 0.01), \quad oX_2 = Bin(3, 0.4)
\]

\[
\beta = (\beta_0 \quad 1.7 \quad -2.5 \quad -2.5)^T, \quad \beta_0 = \{-1.8, -2.0, -2.2\}, \quad \rho = \{0.4, 0.5, 0.6\}
\]

\[
cov(uX_1, oX_1) = -0.2
\]
Evaluating estimator performance: Outcome measure

Weighted average outcome $\hat{y}$

$$y_i \sim N(\mu_y, \sigma_y),$$

$$\mu_y = 3000 + \left\{\frac{300 \times uX_1}{\text{sp.cor.}}\right\} - \left\{\frac{950 \times oX_1 + 200 \times oX_2}{\text{sp.cor.}}\right\},$$

$$\sigma_y = 300$$

Look at the MSE of $\hat{y}$:

$$MSE_y = \mathbb{E} \left[ (\hat{y} - \bar{Y})^2 \right] = \mathbb{E} \left[ (\hat{y} - \mathbb{E}[\hat{y}])^2 \right] + (\mathbb{E}[\hat{y}] - \bar{Y})^2$$
Main Scenarios

- 3 values of $\rho$
- 2 different geographic resolutions (25 vs. 100 subgrids)
- 6 different types of models (GLM, SEM, SAR, GLMM, SEM with aggregated characteristic, SAR with aggregated characteristic)
- 1000 observations
- 1000 replications

$\Rightarrow 3 \times 2 \times 6 \times 1000 = 36'000$ models ($\times 3 \times 3$ for different $cov$ matrix and different response prob.)
Robustness

- ✓ Change response rate:
  40 vs. 50 vs. 60 percent
- ✓ Change covariance structure among independent variables:
  off-diagonal elements -0.2 vs. -0.35 vs. -0.5
- ✓ Varying amount of weight-trimming:
  max. weight 99% vs. 95% vs. 90% quantile
- ✗ Misspecify weights matrix
What does the «world» look like?

Figure 2: Average $uX1$ in the artificial world
1. Spatial Dependence

Figure 3: Moran’s I: Response
Figure 4: Moran’s I: Outcome
2. Model Performance

Figure 5: Share correctly classified $\mathbb{E}[\phi(X)] \approx 0.4$, $\text{cov}(uX, oX) = -0.2$, Pseudo $R^2 \approx 0.25$
Figure 6: Moran’s I: GLM Residuals
Figure 7: Moran’s I: GLMM Residuals
3.1. Weight trimming

Figure 8: MSE trimmed weighted average outcome $E[\phi(X)] \approx 0.4$, $cov(uX1, oX1) = -0.2$, max.

weight: 99% quantile
### 3.1. Weight trimming

Figure 8: MSE trimmed weighted average outcome $\mathbb{E}[\phi(X)] \approx 0.4$, $\text{cov}(uX_1, oX_1) = -0.2$, max. weight: 95% quantile
3.1. Weight trimming

Figure 8: MSE trimmed weighted average outcome $\mathbb{E}[\phi(\mathbf{X})] \approx 0.4$, $\text{cov}(uX_1, oX_1) = -0.2$, max. weight: 90% quantile
Conclusion

- Nonresponse is often geographically dependent
  - due to unobserved selection processes ...
  - ... rather than a contagious behavior
  - Nevertheless: Treat as if contagious to incorporate others’ response status

- Incorporating response of neighboring units increases accuracy of $\hat{\phi}_i$, even in the presence of omitted variable bias

- Especially if large weights are trimmed

- However:
  - Advantages only if spatial models pick up residual spatial correlation after accounting for higher level variance
  - Depends on underlying spatial association, response propensity, and covariance structure
Open topics

- Misspecified weights matrix
- Association of outcome measure with spatial process
- Real-world benchmark
- Vary power of prediction model
Thanks for the attention!


Bibliography II


Appendix
Idea

Make use of spatial dependence in non-response.
Two models to consider:

\[ Y^* = \rho W Y + X \beta + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I) \]  
⇔ \[ Y^* = (I - \rho W)^{-1} (X \beta + \varepsilon) \]  
Spatial Lag Model (SAR)

\[ Y^* = X \beta + \nu, \quad \nu = \lambda W \nu + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I) \]  
⇔ \[ Y^* = X \beta + (I - \gamma W)^{-1} \varepsilon \]  
Spatial Error Model (SEM)
What does the «world» look like?

Figure A.1: Example distribution of response in simulated data
3.2. Response rate

Figure A.2: MSE trimmed weighted average outcome $\mathbb{E}[\phi(X)] \approx 0.4$, $\text{cov}(uX_1, oX_1) = -0.2$, max. weight: 99% quantile
3.2. Response rate

Figure A.2: MSE trimmed weighted average outcome $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = -0.2$, max. weight: 99% quantile
3.2. Response rate

Figure A.2: MSE trimmed weighted average outcome $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX1, oX1) = -0.2$, max. weight: 99% quantile
3.3. Covariance Structure

Figure A.3: MSE trimmed weighted average outcome $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = 0.2$, max. weight: 99% quantile
3.3. Covariance Structure

\[
\rho = 0.4 \quad \rho = 0.5 \quad \rho = 0.6
\]

\begin{align*}
25 & \quad 100 & \quad 25 & \quad 100 & \quad 25 & \quad 100 \\
0 & \quad 30000 & \quad 60000 & \quad 90000 \\
\end{align*}

Number of clusters

MSE estimated mean outcome variable

Estimator
GLM
SEM
SAR
GLMM
SEM agg
SAR agg

Figure A.3: MSE trimmed weighted average outcome $E[\phi(X)] \approx 0.5$, $cov(uX_1, oX_1) = -0.35$, max. weight: 99% quantile
3.3. Covariance Structure

Figure A.3: MSE trimmed weighted average outcome $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = -0.50$, max. weight: 99% quantile
Spatial Dependence
Figure A.4: Moran’s I Response $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = 0.20$
Figure A.5: Moran’s I Response $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = -0.35$
Figure A.6: Moran’s I Response $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX1, oX1) = -0.50$
Figure A.7: Moran’s I Outcome $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX1, oX1) = 0.20$
Figure A.8: Moran’s I Outcome $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = -0.35$
Figure A.9: Moran’s I Outcome $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = -0.50$
Figure A.10: Moran’s I GLM Residuals $E[\phi(X)] \approx 0.5$, $\text{cov}(uX1, oX1) = 0.20$
Figure A.11: Moran’s I GLM Residuals $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = -0.35$
Figure A.12: Moran’s I GLM Residuals $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = -0.50$
Figure A.13: Moran’s I GLMM Residuals $E[\phi(X)] \approx 0.5$, $cov(uX_1, oX_1) = 0.20$
Figure A.14: Moran’s I GLMM Residuals $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = -0.35$
Figure A.15: Moran’s I GLMM Residuals $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX_1, oX_1) = -0.50$
Figure A.16: Moran’s I Response $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = 0.20$
Figure A.17: Moran’s I Response $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = -0.35$
Figure A.18: Moran’s I Response $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = -0.50$
Figure A.19: Moran’s I Outcome $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = 0.20$
Figure A.20: Moran’s I Outcome $\mathbb{E}[\phi(X)] \approx 0.6$, $cov(uX_1, oX_1) = -0.35$
Figure A.21: Moran’s I Outcome $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = -0.50$
Figure A.22: Moran’s I GLM Residuals $E[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = 0.20$
Figure A.23: Moran’s I GLM Residuals $\mathbb{E}[\phi(X)] \approx 0.6$, $cov(uX_1, oX_1) = -0.35$
Figure A.24: Moran’s I GLM Residuals $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = -0.50$
Figure A.25: Moran’s I GLMM Residuals $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = 0.20$
Figure A.26: Moran’s I GLMM Residuals $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = -0.35$
Figure A.27: Moran’s I GLMM Residuals $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = -0.50$
Model performance
Figure A.28: Share correctly classified $\mathbb{E}[\phi(X)] \approx 0.5$, $cov(uX_1, oX_1) = 0.20$
Figure A.29: Share correctly classified $\mathbb{E}[\phi(X)] \approx 0.5$, $\text{cov}(uX1, oX1) = -0.35$
Figure A.30: Share correctly classified $\mathbb{E}[\phi(X)] \approx 0.5$, $cov(uX_1, oX_1) = -0.50$
Figure A.31: Share correctly classified $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = 0.20$
Figure A.32: Share correctly classified $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX1, oX1) = -0.35$
Figure A.33: Share correctly classified $\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX, oX) = -0.50$
MSE weighted average (with weight trimming)
Figure A.34: Share correctly classified

\(\mathbb{E}[\phi(X)] \approx 0.5, \text{cov}(uX1, oX1) = 0.20, \text{max.weight} : 99\% \text{ quantile}\)
Figure A.35: Share correctly classified

\[ \mathbb{E}[\phi(X)] \approx 0.5, \quad \text{cov}(uX_1, oX_1) = -0.35, \quad \text{max.weight} : 99\% \text{ quantile} \]
Figure A.36: Share correctly classified

\[ \mathbb{E}[\phi(X)] \approx 0.5, \quad \text{cov}(uX_1, oX_1) = -0.50, \quad \text{max.weight} : 99\% \text{ quantile} \]
Figure A.37: Share correctly classified

\[ \mathbb{E}[\phi(X)] \approx 0.5, \quad \text{cov}(uX_1, oX_1) = 0.20, \quad \text{max.weight} : 95\% \text{ quantile} \]
Figure A.38: Share correctly classified

$$E[\phi(X)] \approx 0.5, \text{ cov}(uX_1, oX_1) = -0.35, \text{ max.weight} : 95\% \text{ quantile}$$
Figure A.39: Share correctly classified

\[ \mathbb{E}[\phi(X)] \approx 0.5, \quad \text{cov}(uX_1, oX_1) = -0.50, \quad \text{max.weight} : 95\% \text{ quantile} \]
Figure A.40: Share correctly classified
\[ \mathbb{E}[\phi(X)] \approx 0.5, \ covariance(uX_1, oX_1) = 0.20, \text{max.weight : 90\% quantile} \]
Figure A.41: Share correctly classified
\[ \mathbb{E}[\phi(X)] \approx 0.5, \quad \text{cov}(uX_1, oX_1) = -0.35, \quad \text{max.weight : 90\% quantile} \]
Figure A.42: Share correctly classified

\[ E[\phi(X)] \approx 0.5, \quad \text{cov}(uX_1, oX_1) = -0.50, \quad \text{max.weight} : 90\% \text{ quantile} \]
Figure A.43: Share correctly classified

\[ E[\phi(X)] \approx 0.6, \quad \text{cov}(uX_1, oX_1) = 0.20, \quad \text{max.weight} : 99\% \text{ quantile} \]
Figure A.44: Share correctly classified. 
$E[\phi(X)] \approx 0.6$, $cov(uX_1, oX_1) = -0.35$, *max.weight*: 99% quantile.
Figure A.45: Share correctly classified

\[ \mathbb{E}[\phi(X)] \approx 0.6, \quad \text{cov}(uX_1, oX_1) = -0.50, \quad \text{max. weight: 99\% quantile} \]
Figure A.46: Share correctly classified
\[ E[\phi(X)] \approx 0.6, \; cov(uX_1, oX_1) = 0.20, max.weight : 95\% \text{ quantile} \]
Figure A.47: Share correctly classified

$E[\phi(X)] \approx 0.6$, $cov(uX1, oX1) = -0.35$, $max.weight : 95\%$ quantile
Figure A.48: Share correctly classified

$\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = -0.50, \text{max.weight} : 95\%$ quantile
Figure A.49: Share correctly classified

\[ E[\phi(X)] \approx 0.6, \quad \text{cov}(uX_1, oX_1) = 0.20, \quad \text{max.weight : 90\% quantile} \]
Figure A.50: Share correctly classified
\[ E[\phi(X)] \approx 0.6, \; cov(uX_1, oX_1) = -0.35, \; \text{max.weight : 90\% quantile} \]
$\mathbb{E}[\phi(X)] \approx 0.6$, $\text{cov}(uX_1, oX_1) = -0.50$, max.weight : 90% quantile