Online Appendix for "What Do Lost Wallets Tell Us About Survey Measures of Social Capital?"

David Tannenbaum ⓡ Alain Cohn ⓡ Christian Lukas Zünd ⓡ Michel André Maréchal

June 2020

Contents

1 Description of Variables

Generalized Trust Country-level average based on responses to the following question in the World Value Survey (WVS)[1]: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" The variable is recoded as (0) "Need to be very careful" and (1) "Most people can be trusted", so that higher values correspond to higher levels of trust. Standardized at the country-level to have a mean of zero and standard deviation of one.

Generalized Morality Following Tabellini (2008), we compute the fraction of respondents in the World Value Survey (WVS) who select "tolerance and respect for other people" as one of their answers to the question "Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important? Please choose up to five!" Standardized at the country-level to have a mean of zero and standard deviation of one.

Universal Moral Values Following Enke (2019), we computed the difference between the strength of "universal" and "communal" moral values in the Moral Foundations Questionnaire (Graham et al. 2011). Universal moral values are measured as the sum of responses to all questions in the "fairness/reciprocity" and "harm/care" domains, while communal values are measured as the responses to questions in the "in-group/loyalty" and "authority/respect" domains. Standardized at the country-level to have a mean of zero and standard deviation of one.

Civic Cooperation Following Guiso, Sapienza, and Zingales (2011), we computed the first principal component (extracted at the country-level) from the following questions in the World Value Survey (WVS): "Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this card." (1) "Claiming government benefits to which you are not entitled", (2) "Avoiding a fare on public transport", and (3) "Someone accepting a bribe in the course of their duties" on a 10-point scale (from 0 "never justifiable" to 10 "always justifiable"). The principal component has a mean of zero and standard deviation of one, and is coded so that higher values correspond to stronger disapproval.

Trust (GPS) Country-level average based on responses to the following question from the Global Preference Survey (Falk et al. 2018): "I assume that people have only the best intentions" (from 0 "does not describe me at all" to 10 "describes me perfectly"). Standardized at the country-level to have a mean of zero and standard deviation of one.

Positive Reciprocity (GPS) Country-level average based on responses to the following questions from the Global Preference Survey (Falk et al. 2018): (1) "When someone does me a favor I am

1. Not all questions and countries are included in each wave of the World Value Survey. For each country, we use the most recent wave (up to WVS6) in which the question was asked.

willing to return it" and (2) "Please think about what you would do in the following situation. You are in an area you are not familiar with, and you realize you lost your way. You ask a stranger for directions. The stranger offers to take you to your destination. Helping you costs the stranger about 20 Euro in total. However, the stranger says he or she does not want any money from you. You have six presents with you. The cheapest present costs 5 Euro, the most expensive one costs 30 Euro. Do you give one of the presents to the stranger as a 'thank-you'-gift? If so, which present do you give to the stranger?" Falk et al. (2018) aggregate these survey items by computing the z-scores of each question at the individual-level and weighing these z-scores using the weights (0.485 and 0.515, respectively) from the experimental validation procedure in Falk et al. (2016). To facilitate its interpretation, we standardized the measure at the country-level to have mean of zero and standard deviation of one.

Altruism (GPS) Country-level average based on responses to the following questions from the Global Preference Survey (Falk et al. 2018): (1) "How willing are you to give to good causes without expecting anything in return?" and (2) "Imagine the following situation: Today you unexpectedly received 1,000 Euro. How much of this amount would you donate to a good cause?" Falk et al. (2018) aggregate these survey items by computing the z-scores of each question at the individual-level and weighing these z-scores using the weights (0.365 and 0.635, respectively) from the experimental validation procedure in Falk et al. (2016). To facilitate its interpretation, we standardized the measure at the country-level to have mean of zero and standard deviation of one.

Group Membership Following Knack and Keefer (1997), we computed the average number of voluntary social groups that citizens reported belonging to in the World Value Survey (WVS). Respondents are asked whether they belong to each of the following ten groups: (1) "Social welfare services for elderly, handicapped or deprived people", (2) "Religious or church organizations", (3) "Education, arts, music or cultural activities", (4) "Labor unions", (5) "Political parties or groups", (6) "Local community action on issues like poverty, employment, housing, racial equality", (7) "Third world development or human rights", (8) "Conservation, environment, animal rights groups", (9) "Professional associations", (10) "Youth work (scouts, guides, youth clubs, etc.)." Standardized at the country-level to have a mean of zero and standard deviation of one.

Dishonesty Index US The first principal component from (1) the share of self-employed individuals in a city who reports an income in 2009 within US $500 of the first Earned Income Tax Credit (EITC) kink, as a percentage of individuals with non-zero self-employment income, as a measure of cheating on taxes (Chetty, Friedman, and Saez 2013), and (2) the number of federal court convictions for corrupt practices between 1976 and 2002 per 10,000 public officials in each US state (Glaeser and Saks 2006). Following standard practice, we exclude Washington, D.C. from the analysis as the presence of the federal government makes a meaningful comparison with other states difficult (Saiz and Simonsohn 2013). The principal component has a mean of zero and standard deviation of one, and is coded so that higher values correspond to higher levels of dishonest behavior.

Dishonesty Index Italy The first principal component from (1) municipality-level rates of compliance or payment of a television licensing fee (Buonanno et al. 2019), (2) the difference between the cumulative amounts of public money allocated to capital expenditures and existing amounts of physical infrastructure (Golden and Picci 2006), and (3) historical data (1948-1994) on prosecutors' requests to proceed with a criminal investigation against a member of parliament from the city's electoral district (Nannicini et al. 2013). The principal component has a mean of zero and standard deviation of one, and is coded so that higher values correspond to higher levels of dishonest behavior.

Log GDP per Capita The logarithm of countrys' gross domestic product per capita in 2017. Based on the real gross domestic product at constant national prices (in mil. 2011 US $) and the population (in mil.) from the Penn World Table 9.1.

Log Productivity (TFP) The logarithm of a countrys' total factor productivity at current purchasing power parities in 2017 from the Penn World Table 9.1. The variable measures the ratio of a country's economic outputs to its inputs (capital and labor), and this ratio is scaled relative to the United States (which takes a value of 1). For details of the construction, see Feenstra, Inklaar, and Timmer (2015).

Government Effectiveness The estimate of government effectiveness in 2017 from the World Bank (Kraay, Kaufmann, and Mastruzzi 2010). Government effectiveness measures the quality of public services, the quality and degree of independence from political pressures of the civil service, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. Estimates are standardized to have a mean of zero and standard deviation of one.

Letter Grade Efficiency The fraction of correctly returned non-deliverable letters. Chong et al. (2014) mailed ten letters to non-existing business addresses in each country and measured whether they arrived back to a return address in the United States.

Trust in Strangers (online appendix only) Country-level average based on responses to the following question from the World Value Survey (WVS): "Could you tell me for each whether you trust people from this group completely, somewhat, not very much or not at all? People you meet for the first time." The variable is recoded as (1) "Do not trust at all", (2) "Do not trust very much", (3) "Trust somewhat", (4) "Trust completely", so that higher values correspond to higher levels of trust. Standardized at the country-level to have a mean of zero and standard deviation of one.

2 External Validation of Wallet Reporting Rates

We examine variation in wallet reporting rates within the United States and Italy. We focus on these two countries because we sampled a greater number of cities in these countries (e.g., 25 cities in the U.S. compared to 4-5 in most countries), and because established city-level behavior-based proxies of dishonesty/corruption are available.

U.S. Analysis We compared city-level wallet reporting rates to two measures of corruption. Our dishonesty index was constructed by extracting the first principal component from (1) the share of self-employed individuals in a city who reported an income in 2009 within US $500 of the first Earned Income Tax Credit (EITC) kink relative to individuals with non-zero self-employment income, as a measure of cheating on taxes (Chetty, Friedman, and Saez 2013), and (2) the number of federal court convictions for corrupt practices between 1976 and 2002 per 10,000 public officials in the state that the city belongs to (Glaeser and Saks 2006). We use an identical regression specification to that in the main text, except that we restrict ourselves to U.S. data and cluster standard errors at the city-level.

Table A1 reports the relationship between wallet reporting rates and dishonest behavior. We find that a one standard deviation increase in our U.S. dishonesty index is associated with a 4.4 percentage point decrease in reporting a lost wallet ($p = 0.016$). Columns 2 and 3 illustrate that this pattern holds when examining each proxy measure of dishonest behavior separately (for these analyses we standardized each variable to have a mean of zero and SD of one, to facilitate comparison across models).

Italy Analysis We compared city-level wallet reporting rates to three measures of dishonesty. Our dishonesty index was constructed by extracting the first principal component from (1) municipality-level rates of compliance or payment of a television licensing fee (Buonanno et al. 2019), (2) the difference between the cumulative amounts of public money allocated to capital expenditures and existing amounts of physical infrastructure (Golden and Picci 2006), and (3) prosecutors' requests to proceed with a criminal investigation against members of Parliament (Nannicini et al. 2013). We use an identical regression specification to that in the main text, except that we restrict ourselves to data from Italy and cluster standard errors at the city-level.

Table A2 reports the relationship between wallet reporting rates and dishonest behavior. We find that a one standard deviation increase in our Italy dishonesty index is associated with a 7.0 percentage point decrease in reporting a lost wallet ($p = 0.001$). Columns 2-4 illustrate that this pattern holds when examining each proxy measure of dishonest behavior separately (for these analyses we standardized each variable to have a mean of zero and standard deviation of one, to facilitate comparison across models).

Table A1: Wallet Reporting Rates and Dishonesty (USA)

	(1)	(2)	(3)
Dishonesty Index	−4.403		
	(1.686)		
Cheating on Taxes		−3.446	
		(1.541)	
Corruption Convictions			−4.404
			(1.680)
Controls:			
Institution FE	yes	yes	yes
Recipient FE	yes	yes	yes
Situation FE	yes	yes	yes
Treatment FE	yes	yes	yes
Observations	970	1000	970
Cities	24	25	24

Notes: OLS estimates with city-clustered standard errors in parentheses. The dependent variable in all models takes a value of 100 if a wallet was reported and 0 otherwise. "Dishonesty index" is the first principal component from (1) "cheating on taxes": the share of self-employed individuals in a city who reports an income in 2009 within US $500 of the first Earned Income Tax Credit (EITC) kink, as a percentage of individuals with non-zero self-employment income, as a measure of cheating on taxes (Chetty, Friedman, and Saez 2013), and (2) "corruption convictions": the number of federal court convictions for corrupt practices between 1976 and 2002 per 10,000 public officials in each US state (Glaeser and Saks 2006). All explanatory variables are standardized to have a mean of zero and standard deviation of one. All models include controls for the type of institution the wallet drop-off was performed at, characteristics about the recipient of the lost wallet (gender, age), situational characteristics of the wallet drop-off (the presence of a computer, coworkers, or other bystanders), and treatment condition. For full details on control variables see Cohn et al. (2019).

Table A2: Wallet Reporting Rates and Dishonesty (Italy)

	(1)	(2)	(3)	(4)
Dishonesty Index	−7.018			
	(1.593)			
TV Taxes		−3.663		
		(1.723)		
RAPs per Deputy			−7.171	
			(1.115)	
Golden-Picci score				−5.567
				(2.124)
Controls:				
Institution FE	yes	yes	yes	yes
Recipient FE	yes	yes	yes	yes
Situation FE	yes	yes	yes	yes
Treatment FE	yes	yes	yes	yes
Observations	400	400	400	400
Cities	16	16	16	16

Notes: OLS estimates with city-clustered standard errors in parentheses. The dependent variable in all models takes a value of 100 if a wallet was reported and 0 otherwise. "Dishonesty index" is the first principal component from (1) "tv taxes": municipality-level rates of failure to pay a television licensing fee (Buonanno et al. 2019), (2) "RAPs per deputy": historical data on prosecutors' requests to proceed with a criminal investigation against a member of parliament from the city's electoral district (Nannicini et al. 2013), and (3) "Golden-Picci score": the difference between the cumulative amounts of public money allocated to capital expenditures and existing amounts of physical infrastructure (Golden and Picci 2006). All explanatory variables are standardized to have a mean of zero and standard deviation of one. All models include controls for the type of institution the wallet drop-off was performed at, characteristics about the recipient of the lost wallet (gender, age), situational characteristics of the wallet drop-off (the presence of a computer, coworkers, or other bystanders), and treatment condition. For full details on control variables see Cohn et al. (2019).

3 Alternative Specifications

In the main text we report OLS estimates. Figure A1 compares these estimates to estimates of the average marginal effect when running a probit model (to facilitate comparisons, the marginal effects are multiplied by 100 to match the scale of the OLS model). As the figure illustrates, the OLS and probit models return virtually identical results.

The OLS regressions in the main text also include an extensive number of controls, including controls for wallet recipient characteristics (e.g., gender, age), situational characteristics (e.g., the presence of a computer, coworkers, or other bystanders), institutional characteristics (e.g., whether the wallet was dropped off at a bank, cultural establishment, post office, hotel, or public office), and treatment condition (NoMoney vs Money). Figure A2 compares our estimates with controls to a baseline model that excludes these controls. As the figure illustrates, including controls in the model tends to modestly decrease coefficient estimates and does not change the rank-ordering of our social capital measures in predicting wallet reporting rates.

Figure 1: OLS vs. Probit Estimates

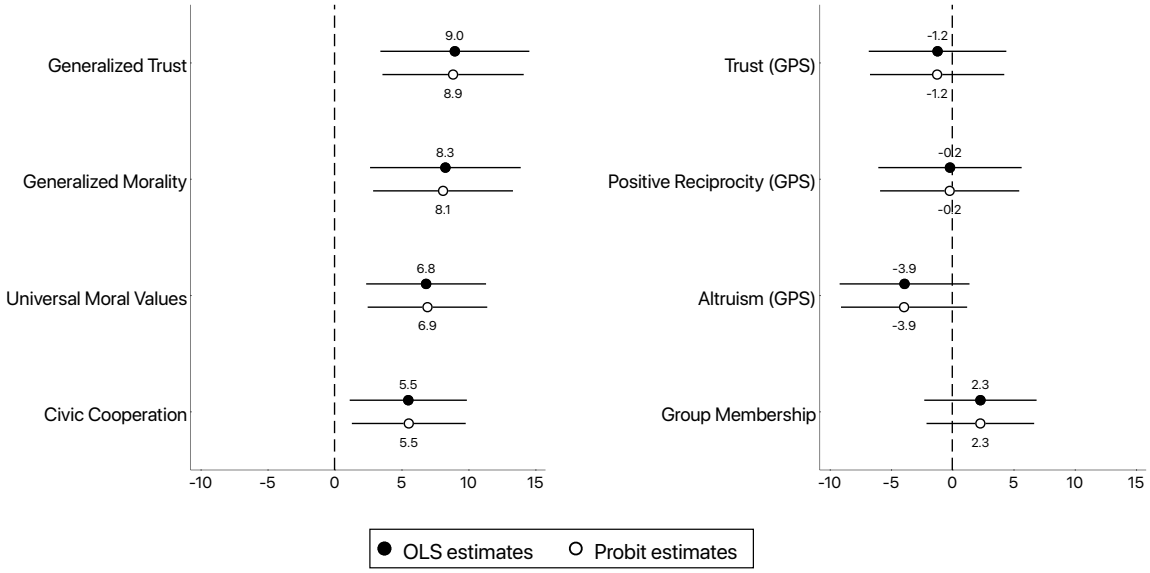
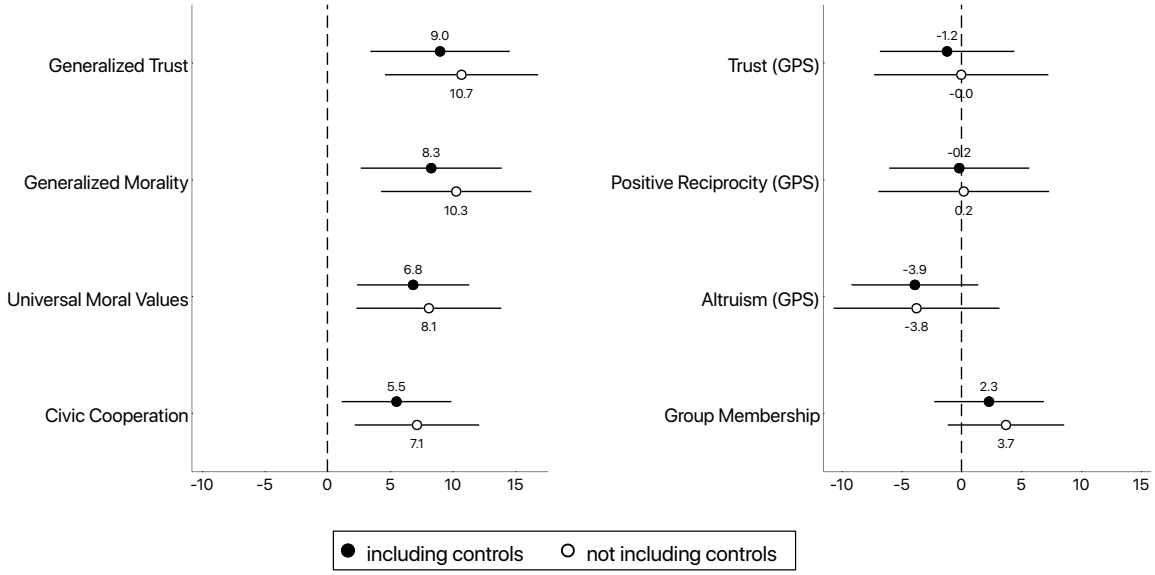


Figure 2: OLS Estimates With Controls vs. Without Controls



4 Correcting for Measurement Error

We use a technique suggested by Gillen, Snowberg, and Yariv (2019) to correct for the role of measurement error. This requires separate and near-identical measurement of the dependent variable (wallet reporting rates) and predictor variable (generalized trust). For each variable we conduct an "obviously related" instrumental variables regression (ORIV) with the replicant instrumented onto the original measurement. The resulting estimate corrects for measurement error. We then adjust the country-level correlation between wallet reporting rates and generalized trust based on these unbiased estimates.

We restrict our analysis to countries in which we had measures of both generalized trust and trust in strangers ($n = 30$). Generalized trust (our predictor variable) was highly correlated with trust in strangers — the country-level rank-order correlated is 0.700 — so we instrumented one variable onto the other. For wallet reporting rates (our dependent variable), recipients in all countries were randomly assigned to either a wallet containing no money or some amount of money. Among the 30 countries in our retained data set the rank-order correlation between the No-Money and Money conditions is 0.933, so we instrumented reporting rates in one condition onto the other condition. As suggested by Gillen, Snowberg, and Yariv (2019), we perform a stacked regression in which each variable serves as both an instrument and an instrumented variable, and then splits the difference between the two estimates. We then adjust the country-level correlation based on these estimates.

For purposes of simplicity, in the main text we report the corrected correlation without adjusting for baseline covariates in the ORIVs. The results are virtually identical when we adjust for baseline covariates: the country-level correlation between generalized trust and wallet reporting rates increases from 0.53 (without measurement error correction) to 0.62 (with measurement error correction).

5 Dominance Analysis

Table A 3 shows the results from dominance analyses for all multivariate models reported in Tables 2 and 3. Dominance analysis (Azen and Budescu 2003; Budescu 1993) is an algorithmic approach to determining the relative contribution of predictors in explaining the variance captured by a regression model. For each multivariate model, the algorithm performs all pairwise comparisons with and without the inclusion of a predictor and calculates the average marginal improvement in the R-squared when that predictor is included in the model. This statistic is then normalized so that the sum of all predictors adds up to 1. For instance, Table 2 in the main text reports an R-squared of 0.673 when generalized trust and wallet reporting rates are used to predict GDP per capita. The results reported in Table A3 indicate that wallet reporting rates contribute 67.4% to that explained variation, and generalized trust contributes 32.6%.

Table A3: Dominance Analysis

	Log GDP (1)	Log TFP (2)	Government Effectiveness (3)	Letter Grade (4)
Wallets	67.4	84.8	53.1	85.9
Generalized Trust	32.6	15.2	46.9	14.1
Wallets	61.2	47.7	51.7	79.7
Generalized Morality	38.8	52.3	48.3	20.3
Wallets	86.9	61.0	94.1	61.4
Universal Moral Values	13.1	39.0	05.9	38.6
Wallets	84.8	65.7	87.3	80.3
Civic Cooperation	15.2	34.3	12.7	19.7
Wallets	69.2	92.3	69.7	98.2
Trust (GPS)	30.8	07.7	30.3	1.8
Wallets	95.1	98.3	97.2	99.7
Positive Reciprocity (GPS)	4.9	1.7	2.8	0.3
Wallets	94.6	98.9	95.5	96.4
Altruism (GPS)	05.4	1.1	4.5	3.6
Wallets	85.5	82.9	68.6	72.9
Group Membership	14.5	17.1	31.4	27.1

Notes: The table reports the results of a dominance analysis between each explanatory variable and wallet reporting rates for (1) log GDP per capita, and (2) log total factor productivity (relative to the United States), (3) government effectiveness ratings from the World Bank, and (4) letter grade efficiency scores from Chong et al. (2014). For each pair of explanatory variables, numbers represent the percentage contribution of that variable to the total R^2 (with pairs adding up to 100%).

References

Azen, R, and DV Budescu. 2003. "The dominance analysis approach for comparing predictors in multiple regression." *Psychological Methods* 8 (2): 129–148.

Budescu, David. 1993. "Dominance Analysis: A New Approach to the Problem of Relative Importance of Predictors in Multiple Regression." *Psychological Bulletin* 114 (3): 542–551.

Buonanno, Paolo, Matteo Cervellati, Sara Lazzaroni, and Giovanni Prarolo. 2019. "Political History, Fiscal Compliance and Cooperation: Medieval Social Contracts and Their Legacy."

Chetty, Raj, John N Friedman, and Emmanuel Saez. 2013. "Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings." *American Economic Review* 103 (7): 2683–2721.

Chong, Alberto, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2014. "Letter grading government efficiency." *Journal of the European Economic Association* 12 (2): 277–298.

Enke, Benjamin. 2019. "Kinship, cooperation, and the evolution of moral systems." *The Quarterly Journal of Economics* 134 (2): 953–1019.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global evidence on economic preferences." *The Quarterly Journal of Economics* 133 (4): 1645–1692.

Falk, Armin, Anke Becker, Thomas J Dohmen, David Huffman, and Uwe Sunde. 2016. "The preference survey module: A validated instrument for measuring risk, time, and social preferences." Netspar Discussion Paper No. 01/2016-003. Available at SSRN: https://ssrn.com/abstract=2725874.

Feenstra, Robert C, Robert Inklaar, and Marcel P Timmer. 2015. "The next generation of the Penn World Table." *American economic review* 105 (10): 3150–82.

Gillen, Ben, Erik Snowberg, and Leeat Yariv. 2019. "Experimenting with measurement error: Techniques with applications to the caltech cohort study." *Journal of Political Economy* 127 (4): 1826–1863.

Glaeser, Edward L, and Raven E Saks. 2006. "Corruption in America." *Journal of Public Economics* 6 (90): 1053–1072.

Golden, Miriam, and Lucio Picci. 2006. "Corruption and the management of public works in Italy." In *International Handbook on the Economics of Corruption,* 457.

Graham, Jesse, Brian Nosek, Jonathan Haidt, Ravi Iyer, Spassena Koleva, and Peter Ditto. 2011. "Mapping the Moral Domain." *Journal of Personality and Social Psychology* 101 (2): 366–385.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2011. "Civic capital as the missing link." In *Handbook of Social Economics,* edited by J. Benhabib, A. Bisin, and M. O. Jackson, 1:417–480. Amsterdam: North-Holland.

Knack, Stephen, and Philip Keefer. 1997. "Does social capital have an economic payoff? A cross-country investigation." *Quarterly Journal of Economics* 112 (4): 1251–1288.

Kraay, Aart, Daniel Kaufmann, and Massimo Mastruzzi. 2010. *The Worldwide Governance Indicators: Methodology and Analytical Issues.* The World Bank.

Nannicini, Tommaso, Andrea Stella, Guido Tabellini, and Ugo Troiano. 2013. "Social capital and political accountability." *American Economic Journal: Economic Policy* 5 (2): 222–50.

Saiz, Albert, and Uri Simonsohn. 2013. "Proxying for unobservable variables with internet document-frequency." *Journal of the European Economic Association* 11 (1): 137–165.

Tabellini, Guido. 2008. "Institutions and culture." *Journal of the European Economic Association* 6 (2-3): 255–294.