SUPPLEMENT TO "WHITHER FORMAL CONTRACTS?"

RAÚL SÁNCHEZ DE LA SIERRA
Harris School of Public Policy, University of Chicago and National Bureau of Economic Research

APPENDIX A: TABLES AND FIGURES

TABLE A.I

WHO ARE THE MERCHANTS?[a]

	All Sample
A: Merchants' basic characteristics	
Observations	20
Household size	2.44
Age	30.28
Years of education	13.17
Main occuption: public	0.05
Main occuption: student	0.45
Main occuption: private sector	0.06
Main occuption: NGO	0.33
State-favored: *Bashi*	0.40
State-favored: *Other*	0.15
B: Buyers' beliefs about merchants' main occupation	
Observations	958
Occupation: public	0.11
Believes merchant works for foreigners	0.02
Believes merchant works for government	0.12
Believes merchant is a student	0.38
Believes merchant works for firm	0.05
Believes merchant works for NGO	0.20

[a]Source for Panel A: merchants' survey. Source for Panel B: perceptions survey. All numbers are means of indicator variables except for age, which reports the mean number of years since the individual was born, years of education, which is in years, and household size, which is mean number of individuals.

Raúl Sánchez de la Sierra: raul@uchicago.edu

TABLE A.II

PERCEPTIONS SURVEY RESPONDENTS' BELIEFS[a]

	Mean	N
A: Beliefs about the sale		
A: Beliefs about sale design		
Door to door sales' frequency (days/year)	142.53	439
Sale on credit is normal	0.60	453
B: Beliefs about governance of informal agreements		
Credible threats, of which:	0.64	482
Reputation	0.24	486
Instantaneous sanctions (threats of violence)	0.17	486
State, predation	0.14	486
State, formal prosecution	0.24	486
C: Beliefs about governance of formal agreements		
Believes this contract is normal	0.37	472
Uses this contract habitually	0.35	478
This contract induces compliance, because:	0.84	454
Reputation	0.31	486
Instantaneous sanctions (threats of violence)	0.06	486
State, predation	0.04	486
State, formal prosecution	0.43	486

	Merchants	S	D	D-S (*p*-Value)
B: Beliefs about the merchants shown in the photos				
Observations	20	10	10	
Merchant seen as state-favored	0.64	0.88	0.33	0.00
Merchant seen as trusting	0.40	0.42	0.37	0.19
Merchant seen as trustworthy	0.40	0.41	0.38	0.39
Merchant seen as friendly	0.58	0.61	0.55	0.29
Merchant's I.A.T. score	0.03	0.02	0.04	0.64

[a]Source: perceptions survey. Percentages are the mean of the response for each question. Unless otherwise noted, the responses are indicator variables taking value 1 if the respondent mentioned the reported answer.

TABLE A.III

FORMAL VERSUS INFORMAL TRADE [STATE-FAVORED VS. DISFAVORED MERCHANTS] EXCL. DISFAVORED BUYERS; C: CHEAT (BUY AND DEFECT); E: EXCHANGE (BUY AND PAY); R: REJECT[a]

		Buyers' Choices			Buyers' Beliefs			
	Dependent:	$I_C(b_i)$	$I_E(b_i)$	$I_R(b_i)$	*Rep.*	*Shame.*	*Viol.*	*Pros.*
Explanatory		(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Buyer-merchant pairs with state-favored merchant: lifted vs. maintained								
Requirement maintained, $\tilde{F}_i$		−0.18	0.05	0.13	0.03	0.02	0.01	0.12
		(0.05)	(0.04)	(0.04)	(0.02)	(0.03)	(0.01)	(0.04)
Obs.		399	399	399	399	399	399	399
R^2		0.23	0.19	0.22	0.47	0.28	0.20	0.24
Mean dep.		0.60	0.17	0.23	0.07	0.14	0.08	0.17
Multinomial coeff.			0.79	0.97				
Multinomial SE.			0.31	0.27				
B. Buyer-merchant pairs with disfavored merchant: lifted vs. maintained								
Requirement maintained, $\tilde{F}_i$		0.01	−0.03	0.02	−0.02	0.02	0.01	0.02
		(0.06)	(0.04)	(0.05)	(0.03)	(0.03)	(0.02)	(0.03)
Obs.		291	291	291	291	291	291	291
R^2		0.18	0.12	0.25	0.39	0.29	0.17	0.31
Mean dep.		0.50	0.17	0.31	0.09	0.08	0.08	0.08
Multinomial coeff.			−0.09	0.08				
Multinomial SE.			0.37	0.31				

[a]0. Estimates from equation (1), estimated in samples described in (i) in Section 4.5 ($l_i \in S$ vs. $l_i = D$). In addition, the sample in both panels also excludes all buyers that belong to the disfavored set and those that are state officials.
1. Dependent variables are in panel headers, coefficients in body of table, and standard errors are in parentheses.
2. Dependent variables are indicator variables taking value 1 for the following ultimate choices of the buyer in the sale of credit and zero otherwise, from left to right: buys and defects ($I_C(b_i)$), buys and pays ($I_E(b_i)$), rejects ($I_R(b_i)$), believes that the following sanctions would likely ensue if she defected: reputation loss (rep.), psychological costs (shame), violence (viol.), formal prosecution (pros.). Explanatory variable, "Requirement maintained," is an indicator for whether the requirement to sign the contract is maintained for buyer i, $\tilde{F}_i \in \{0, 1\}$.
3. In Panel A, the sample is all buyers matched to a state-favored merchant ($l_i \in S$).
4. In Panel B, the sample is all buyers matched to a disfavored merchant ($l_i = D$).
5. Mean dep. reports the mean of the dependent variable in the comparison group of the corresponding sample.
6. Multinomial coeff. and Multinomial SE. report the logit coefficients and their respective standard errors in a multinomial logistic regression of buyer choice on the corresponding independent variable, base choice $b = C$.
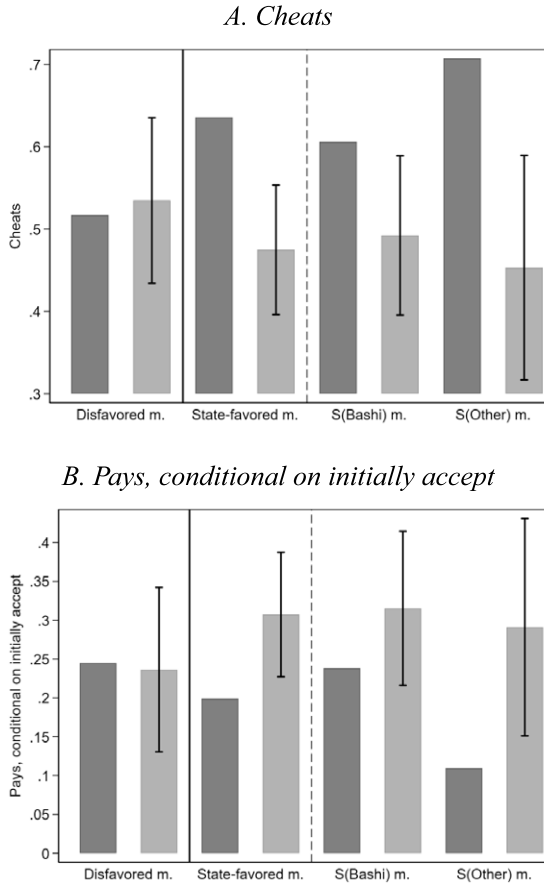
A. Cheats

B. Pays, conditional on initially accept

FIGURE A.1.—Formal contract effect, decomposed by merchant's ethnic group. *Notes*: Dark bar represents buyers assigned to contract lifted. Light bar represents buyers assigned to contract maintained. Panel A: dependent variable is an indicator for whether the buyer chooses to cheat in the sale on credit, $I_C(b_i)$. Panel B: dependent variable is decision to pay in the sale on credit, among buyers who initially accept, $I_E^A(b_i)$. Bracketed lines represent 95% confidence intervals, estimated in corresponding regression. Disfavored m., State-favored m., S(Bashi) m., S(Other) m. indicate that the buyer-merchant match is $l_i = D$, $l_i \in S$, $l_i \in S$ in which the merchant is Bashi, $l_i \in S$ in which the merchant is from another state-favored ethnic group, respectively.
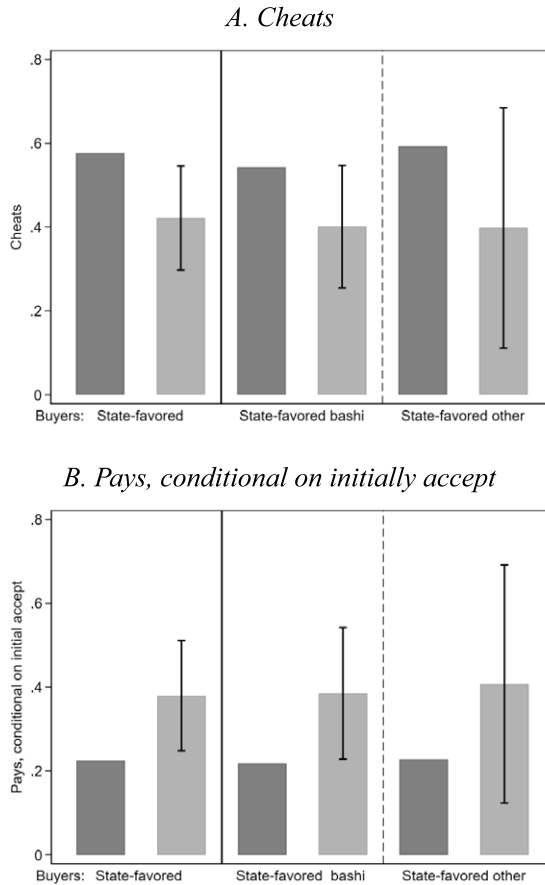
A. Cheats

B. Pays, conditional on initially accept

FIGURE A.2.—Co-ethnicity effect in sale on credit, decomposed, by buyer's ethnic group. *Notes*: Dark bar represents means for buyers in non-co-ethnic pair. Light bar represents means for buyers in co-ethnic pair. Panel A: dependent variable is the decision to cheat in the sale on credit, $I_C(b_i)$. Panel B: dependent variable is decision to pay in the sale on credit, among buyers who initially accept, $I_E^A(b_i)$. Bracketed lines represent 95% confidence intervals, estimated in corresponding regression.

A: *Formal vs. informal trade – state bias and state enforcement with co-ethnic governance*
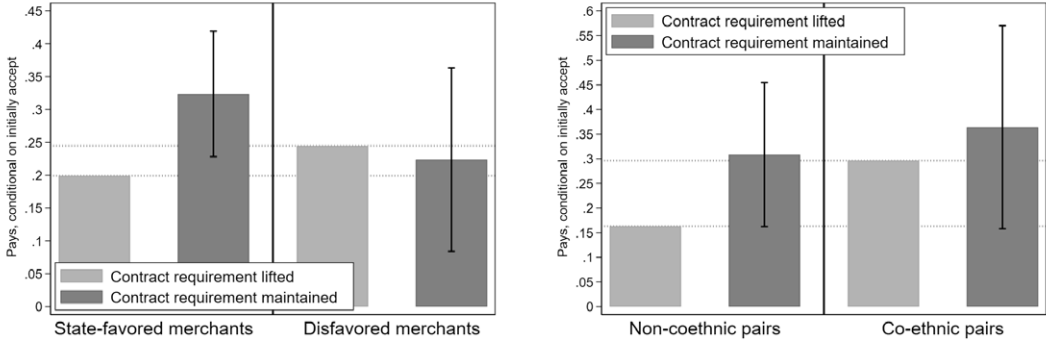
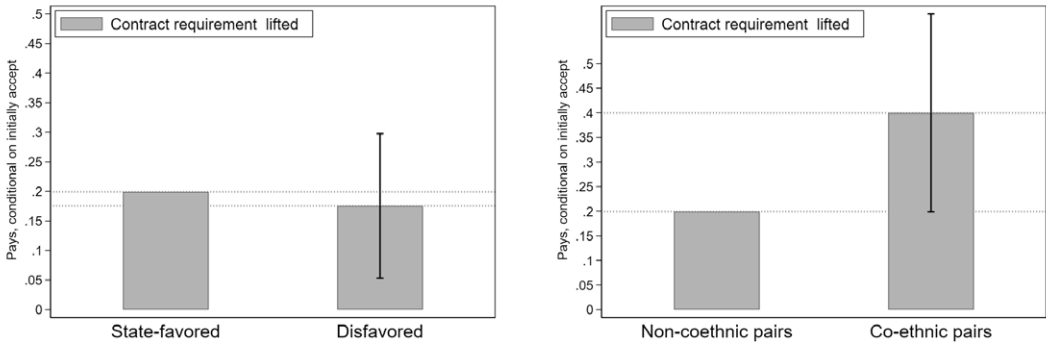**B:** *Informal trade – co-ethnic governance*

FIGURE A.3.—Robustness of analysis excluding selection to controlling for placebo choices. *Notes*: 0. The sample in both panels excludes buyers who initially reject and the dependent variable is whether buyer i pays or defects, $I_E^A(b_i)$. Left quadrants compare buyers matched to a state-favored/disfavored merchant ($l_i \in S$ vs. $l_i = D$), right quadrants compare buyers in non-co-ethnic pairs ($l_i = SN$) to those in co-ethnic pairs ($l_i = SC$). *In addition, I include as control variable an indicator for whether the buyer had rejected the placebo sale, $I_R(b_i^0) \in \{0, 1\}$.* 1. Panel A reports the estimates of equation (1). The explanatory variable is an indicator for whether the buyer is assigned to maintained, $\tilde{F}_i$. In the left panel, the sample is (comparison (i) in description of equation (1)): left quadrant includes all buyers matched to a state-favored merchant ($l_i \in S$), and right quadrant includes all that are matched to a disfavored ($l_i = D$). In the right panel, the sample includes only buyers in state-favored buyer-merchant pairs and is as follows (comparison (ii) in description of equation (1)): left quadrant is buyers in non-co-ethnic pairs ($l_i = SN$), and right quadrant is buyers in co-ethnic pairs ($l_i = SC$). 2. Panel B reports the estimates of equation (2) in the sample of buyers assigned to lifted, $\tilde{F}_i = 0$. In the left panel, the sample is all those buyers ((ii) in description of equation (2)), the explanatory variable is $I_D(l_i) \in \{0, 1\}$. In the right panel, the sample is buyers in state-favored buyer-merchant pairs ((i) in description of equation (2)), the explanatory variable is $I_{SC}(l_i) \in \{0, 1\}$. 3. Significance. Black brackets are 95% confidence intervals on $a^{\tilde{F}}$ (Panel A) and on a^I (Panel B).

TABLE A.IV

THE POWER OF BELIEFS THAT SANCTIONS ARE LIKELY (IV ESTIMATES); C: CHEAT (BUY AND DEFECT); E: EXCHANGE (BUY AND PAY); R: REJECT[a]

	2SLS Estimates			
Dependent Variable:	$I_C(b_i)$	$I_E(b_i)$	$I_R(b_i)$	$I_E^A(b_i)$
Instrumented Variable:	(1)	(2)	(3)	(4)
A. Belief that prosecution is likely (trade with state-favored merchant)				
Believes legal prosecution is likely	−1.27	0.36	0.90	0.65
	(0.47)	(0.25)	(0.39)	(0.29)
Obs.	511	511	511	357
R^2	−1.34	−0.15	−0.92	−0.53
Mean dep.	0.64	0.52	0.30	0.20
First-stage effective F stat.	13.90	13.90	13.90	14.49
Instrumental variable: formal contract				
B. Belief that ethnic sanctions are likely (informal trade with state-favored merchant)				
Believes ethnic sanctions are likely	−1.26	0.48	0.68	0.53
	(0.68)	(0.38)	(0.49)	(0.30)
Obs.	207	207	207	152
R^2	−1.47	−0.43	−0.60	−0.44
Mean dep.	0.69	0.13	0.18	0.21
First-stage effective F stat.	5.51	5.11	5.11	10.44
Instrumental variable: co-ethnic				

[a]1. Two-stage least squares regression, with heteroscedasticity robust standard errors. Panel A: The endogenous regressor in columns (1)–(4) is an indicator taking value 1 if the buyer believes prosecution is likely if she defects, and zero otherwise. The instrumental variable is an indicator for whether the buyer is assigned to contract maintained, $\tilde{F}$, and zero otherwise. Columns (1)–(4) present the regression with the standard controls described in Section 4. Panel B: The endogenous regressor in columns (1)–(4) is an indicator taking value 1 if the buyer believes reputational cost and other informal sanctions relevant for intra-ethnic trade are likely if she defects, and zero otherwise. The instrumental variable is an indicator for whether the buyer is assigned to a co-ethnic pair, $l_i = SC$, among the state-favored buyer-merchant matches, that is, among $l_i \in SC \cup SN$, and zero otherwise. Columns (1)–(4) present the regression with the controls described in Section 4.

2. Dependent variables are in panel headers, coefficients are in body of table, and standard errors are in parentheses.

3. Dependent variables are indicator variables taking value 1 if the buyer rejects the sale on credit and zero otherwise ($I_R(b_i)$), whether she pays and defects and zero otherwise ($I_C(b_i)$), whether she buys and pays and zero otherwise ($I_E(b_i)$), and whether she pays conditional on initially accepting and zero otherwise $I_E^A(b_i)$.

4. Sample. Panel A: all buyers matched to a state-favored merchant. Panel B: all buyers in state-favored buyer-merchant pairs.

5. Mean dep. reports the mean of the dependent variable in the comparison group of the corresponding sample.

6. First-stage effective F stat. reports Olea and Pflueger's (2013) effective F-statistic, robust to heteroscedasticity. An F-statistic larger than 10 indicates a strong instrument (Staiger and Stock (1997)).

7. Results in Panels A and B are identical if controls for other, possibly endogenous, beliefs are included. Specifically, in Panel A, results are preserved if I control for whether the buyer believes that reputation, shame, or violence are likely. In Panel B, results are preserved if I control for whether the buyer believes that legal prosecution, shame, or violence are likely.

8. Comments. In Panel A, I instrument whether the buyer believes that legal prosecution is likely with assignment to maintained. Columns (1), (2), and (3) respectively show that switching the buyer's belief that prosecution is likely from 0 to 1 decreases the probability that the buyer cheats, $I_C(b_i)$, by - 1.27 pp., increases the probability that the buyer buys and pays, $I_E(b_i)$, by 36 pp., and increases the probability that the buyer rejects, $I_R(b_i)$, by 9 pp., all of which are statistically significant effects. The effective F-statistic for the first stage (Olea and Pflueger (2013)) is 13.90, thus the first stage is strong (Staiger and Stock (1997)). Column (4) uses as dependent variable $I_E^A(b_i)$. The conclusion is identical. In Panel B, the endogenous regressor is whether the buyer believes that ethnic sanctions are likely if she defects and the instrument is *Co-ethnic*. The sample is sales assigned to lifted. While the F-statistic in columns (1)–(3) is somewhat weak, columns (1)–(4) show that expecting intra-ethnic sanctions has an effect of similar magnitude to expecting prosecution.

TABLE A.V

ESTIMATION OF BUYERS' MASSES BY MATCH TYPE[a]

	D	$S\{N\}$	$S\{C\}$
		A: Estimands	
α_1:	$\alpha^p(\alpha^b + \beta^b_F)$	$(\alpha^p + \beta^p_S)(\alpha^b + \beta^b_F + \beta^b_s + \beta^b_{FS})$	$(\alpha^p + \beta^p_S + \beta^p_C)(\alpha^b + \beta^b_F + \beta^b_s + \beta^b_{FS} + \beta^b_c + \beta^b_{FC})$
α_2:	$1 - \alpha^b$	$1 - \alpha^b - \beta^b_s$	$1 - \alpha^b - \beta^b_s - \beta^b_c$
α_3:	$\beta^p_F(\alpha^b + \beta^b_F)$	$(\beta^p_F + \beta^p_{FS})(\alpha^b + \beta^b_F + \beta^b_s + \beta^b_{FS})$	$(\beta^p_F + \beta^p_{FS} + \beta^p_{FC})(\alpha^b + \beta^b_F + \beta^b_s + \beta^b_{FS} + \beta^b_c + \beta^b_{FC})$
α_4:	$-\beta^b_F$	$-\beta^b_F - \beta^b_{FS}$	$-\beta^b_F - \beta^b_{FS} - \beta^b_{FC}$
		B: Estimates	
α_1: always peaches	0.17	0.11	0.19
	(0.00)	(0.00)	(0.00)
α_2: never buyers	0.31	0.18	0.30
	(0.00)	(0.00)	(0.00)
α_3: informal lemons, formal peaches	−0.01	0.09	0.05
	(0.72)	(0.03)	(0.35)
α_4: informal lemons, formal non-buyers	0.01	0.16	0.05
	(0.82)	(0.00)	(0.45)
α_5: always lemons	0.52	0.46	0.40
	(0.00)	(0.00)	(0.00)

[a]1. Panel A: Estimands of buyer masses for markets where merchants are disfavored (D), where merchants and buyers are state non-coethnics ($S\{N\}$), where merchants and buyers are state co-ethnics ($S\{C\}$).

2. Panel B: Estimates, derived using coefficient estimates from OLS without randomization block fixed effects. Including randomization block fixed effects yields similar results: α_3 and α_4 for $S\{N\}$ are strengthened and the difference $S\{C\} - S\{N\}$ in row α_1 is weakened. An alternative estimation strategy where the regression of $I_E(b_i)$ is not conditioned on to those that initially accept yields similar results. This alternative strategy does not hinge on the assumption that assignment to formal contract is orthogonal to initially accepting but is less efficient. Using the alternative strategy, the difference $S\{N\} - D$ in row α_3 is weakened but remains statistically significant.

3. Panel B: Columns D through $S\{C\}$. The numbers in parentheses indicate the p-value of a Wald test of the null hypothesis that the estimated combination of parameters is different than zero. I test the hypothesis that $R(z) = q$, where z is a vector of OLS coefficients, $R(z)$ is any nonlinear combination of such parameters, G is the derivative of $R(z)$ with respect to z, and V is a symmetric weighted matrix. I then construct the Wald statistic $W = (R(z) - q)' inv(GVG')(R(z) - q)$. Under the null hypothesis, $R(z) = q$, W follows a χ^2 distribution. In Stata, the statistical test command used is "testnl", which tests nonlinear hypotheses after estimation, because the expressions from α_1 to α_5 are nonlinear combinations of coefficients.

APPENDIX B: ADDITIONAL IMPLEMENTATION DETAILS

This section provides additional details on the implementation of the experiment.

B.1. *Language of the Sale*

The merchant first obtains informed consent (translated from French, original emphasis):

> "Hello, my name is …. I work for a project of scientific research of the University of Columbia in New York. I am not employed by any agency of the state nor by any other group. Our project aims to study the functioning of communities urban areas in eastern Congo to better understand the economic problems from which we suffer and thus make ourselves heard.
>
> In the framework of this university research project, I would be pleased to ask a few questions to your household if it is possible. You have in any case the free choice to participate or to refuse participation in our project. For all the answers that you will give me, I will keep strict anonymity, that is to say that nobody will be able to know what you have *personally* said. If this information is used by us later, no one will know that it came from you personally and your name will not be kept on any table that others can access. Before I ask your agreement to participate in this research, it is advisable to remind you that this research does not include any risk for your NEIGHBORHOOD nor for any individual of this NEIGHBORHOOD. If ever you consider that a response that you will give us could you put in danger, putting in jeopardy someone in your family or in your NEIGHBORHOOD, or I put in danger, it is better not to answer. It should also be noted that research has no direct benefit either. This research aims to improve the quality of our knowledge of the functioning of the urban economy in the east of the DRC and could help the improvement of the quality of the actions international designed to help in the region. If you allow me to proceed, you will help to better understand the situation of the community. I would be very grateful for your help in responding to this questionnaire. The interview does should not last more than 15 Minutes. Would you like your household to participate in this survey? Here is a sheet of information, in which you have of any information contact the event or you would of complaints or the concerns related to this research, or if you want to learn more about your rights as a participant of this research.
>
> Mister/Madam, I have another preoccupation; I am a seller of phone airtime cards. I would like to sell to you a part of my phone airtime cards if you were interested. The sale price I offer includes a reduction of 50% of the price that you would normally pay."

After the placebo sale is complete, the merchant presents the sale on credit (translated from French, original emphasis):

> "Mister/Madam. I also offer you a second card at the same price. However, payment is made by sending telephone units by phone to our office telephone. This procedure is necessary for my supervisor to account for the sale of this card, but part of the money will be my income. Are you interested in purchasing this card? If so, you will have to sign a contract in which you commit to make this payment. Do you accept?"

After the buyer communicates whether they initially accept the offer, the merchant implements the contract randomization (translated from French, original emphasis):

> "Dear Sir Madam. Actually, I don't have enough contracts, so you can get this card, but you have to pay within three days by sending a units by phone. "

The project document then indicates:

> "IMPORTANT: whether it is a DO SIGN household or a FORGIVE household, the CHOICE of accepting or refusing MUST BE MADE immediately after this question:
> 'Are you interested in purchasing this card? If so, you will have to sign a discharge through which you commit legally to make this payment. Do you accept?'
> It is only AFTER observing the choice, that some households will be FORGIVEN the need to sign a waiver. If they signed up with a contract, they have everything to gain when they are taken away from the need to sign the contract, and it is to their advantage.'

Source: project documents of October 24th 2013.
The training documents indicate:

"One of the responsibilities of the enumerators is to adminster the sale of phone airtime cards in the selected households to participate in this research, that is why we strongly encourage the enumerators to be ready to behave like actual vendors in the neighborhood."

Source: training documents of October 7th 2013.

B.2. *Official Documentation*

The study was approved as part of the IRB protocol IRBAAAK0552 (Y1M00), hosted at Columbia University during the author's dissertation. Figures B.1 and B.2 respectively show the city's authorization for all activities, and the contract used.

B.3. *Language of Merchants' Agreement*

The agreement signed by merchants includes the following text (translated from French):

"The finances will consist of 3 components:
– Fixed Salaries of enumerator, which will be paid at the end of the study and after verification of the work of each enumerator.
– Miscellaneous costs: daily transport costs from the house to the study neighborhods. Transport costs will be calculated on the basis of the distance to be covered by each person.
– Income from profit from sales:
[...] enumerators should follow the research instructions and should not go below and/or beyond pre-established instructions. Investigators will buy the cards from Jean–Paul (data collection supervisor) on the basis of a reduced price of 500FC per card and have the recommendation to sell [two cards in each household] in turn at 1000FC reduction per card. All revenues from the sales will be added to the salary."

B.4. *Payment Outcomes*

Buyers committed to transfer funds to the data collection supervisor using the equivalent value in airtime credit (the predecessor to mobile money), and send a message with their household code. Alternatively, buyers could exchange phone numbers with the merchant and meet him within three days to pay cash. If they chose this option, the buyer had to text the confirmation code to the supervisor. Codes and payments were verified by the supervisor and the merchants in daily evening meetings.

B.5. *Additional Checks on Merchant Shirking*

In addition to creating the bound denoted "Adj. coeff" in Table 3, I used the data to investigate whether merchants skipped sale presentation steps. I found that, for just four merchants, initially accept was negatively correlated with assignment to contract. This could arise by chance. It could also reflect that these four merchants shirked. If they did, by coding ultimately accept as initially accept, this could confound the estimate of $\hat{a}^{\tilde{F}}$ in the analysis excluding selection. Three of them were state-favored. Excluding the disfavored had no impact on the estimates of the disfavored analysis. Then, I estimated equation (2) excluding the data from each of the remaining three state merchants, one, then two at a time. The result is unchanged. Excluding all three weakens the coefficient but the magnitude is similar and it remains marginally significant with *p*-value 0.12. Finally, I examined this correlation over time. I found that it is only significant for November 11th, 2013. Thus, I also then estimated equation (2) excluding November 11th. Excluding this

REPUBLIQUE DEMOCRATIQUE DU CONGO
PROVINCE DU SUD-KIVU
MAIRIE DE BUKAVU
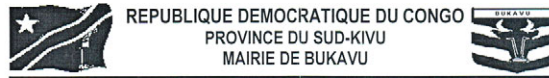
Bukavu, le _03_ / _7_ /2013

N° 401/BUR/M.BKV/ _405_ /2013

Transmis Copie pour Information à :

- Monsieur le CCRG-ANR/Ville,
  à BUKAVU
- Monsieur Raul SANCHEZ de la Sierra,
  Mphil en Economie, Département d'Economie
  Columbia University
  (Tous) aux USA

Objet : *Recherche histoire Economie et Sociale du Sud-Kivu*   A Monsieur Paul ZIBIKA MUTACUMURA,
Coordinateur Provincial pour la
Recherche de l'Histoire Economie
et Sociale du Sud-Kivu
à BUKAVU

Monsieur,

J'ai l'honneur d'accuser réception de la lettre sans numéro du 17 Juin 2013 relative à l'objet sous rubrique m'adressée par Monsieur Raul Sanchez de la Sierra, Mphil en Economie, Département d'Economie à Columbia University.

Conformément à l'esprit de la précitée, je vous autorise de circuler dans la Ville de Bukavu pour mener votre enquête sur l'histoire Economique et Sociale du Sud-Kivu. Toutefois, vous êtes invités au respect strict de l'ordre et la tranquillité publics au cours de cette activité.

Mes sentiments patriotiques.

LE MAIRE DE BUKAVU

Mr. LUTOMBO YOGOLELO Philémon
Directeur

Nos adresses : 137, Bld Patrice Emery Lumumba, Commune d'Ibanda – Site web : www.mairiedebukavu.net
E-mail : info@mairiedebukavu.net

FIGURE B.1.—Authorization for the study activities by the Mayor of Bukavu. *Notes*: Mayor's authorization for the study, obtained after presenting the study protocols to the mayor's office.

day increases the magnitude of $\hat{a}^{\tilde{F}}$ for state-favored merchants and decreases its p-value. Thus, this form of shirking was infrequent and cannot drive the results in the analysis excluding selection. In addition, for robustness I re-coded eight observations with contract requirement maintained in which the choices were coded as initially reject, ultimately accept. Since the merchant recorded an agreed price for these eight observations, and since merchants confirmed that no buyer revisited the choice positively if the contract was maintained, this provides confidence that these eight buyers had initially accepted. The results are unchanged by this correction.

FIGURE B.2.—Contract. *Notes*: Panel A shows the contract that the respond signs when she commits to pay by sending airtime. Panel B shows the receipt of payment given the buyer for those who paid person.

B.6. *Exit Survey Variables not Used in the Paper*

In addition to the variables used in the analysis, the exit survey gathered two additional variables. These are: (a) the axes of solidarity. In this variable, the merchants asked the buyers to list in order the three major dimensions along which in-group solidarity should be invoked for them; (b) to list the main dimension through which the buyer sees the merchant. Both of these questions referred to: ethnic group, nation, and neighborhood. The objective for including these questions was to explore ethnic saliency, but during the project's preparations, it was made clear that the disfavored (Tutsi) ethnic group is associated also with foreignness, and thus this distinction was no longer relevant.

APPENDIX C: BOUNDS WITH STRATEGIC ETHNIC PASSING

Let c and nc be the shares of co-ethnics and non-co-ethnics in the population, and p_c and p_{nc} indicate respectively the share of co-ethnics and of non-co-ethnics in the population who defect. Suppose that, among the defectors who are co-ethnic, share s pass as non-co-ethnics and that the corresponding share for non-defectors, and for non-coethnics, is zero. Let $\tilde{c}$ and $\tilde{nc}$ be the share of buyers reporting to be co-ethnics and non-co-ethnics.

Then, consider the following equation, where $I_E^A(b_i)$ denotes whether a buyer who initially accepts pays: $I_E^A(b_i) = \alpha + aI_{SC}(l_i) + e_i$. We have: $a = E[I_E^A(b_i)|c = 1] - E[I_E^A(b_i)|c = 0]$. Estimating this equation through OLS produces an estimator for a, $\hat{a}$. By definition, $E(\hat{a}) = E[I_E^A(b_i)|\tilde{c} = 1] - E[I_E^A(b_i)|\tilde{c} = 0]$, with the two terms defined as follows:

$$E\big[I_E^A(b_i)|\tilde{c} = 1\big]$$
$$= \frac{c(1 - p_c)}{c\big[(1 - p_c) + p_c(1 - s)\big]} \times 1 + \frac{cp_c(1 - s)}{c\big[(1 - p_c) + p_c(1 - s)\big]} \times 0$$
$$= \frac{1 - p_c}{1 - p_c + p_c(1 - s)}, \tag{3}$$
$$E\big[I_E^A(b_i)|\tilde{c} = 0\big]$$

$$= \frac{nc}{nc + sp_c c} E[I_E^A(b_i)|c=0] + \frac{sp_c c}{nc + sp_c c} \times 0$$

$$= \frac{nc}{nc + sp_c c} E[I_E^A(b_i)|c=0]. \tag{4}$$

We can use equations (3) and (4) to derive an expression for a as follows:

$$E(\hat{a}) = \frac{1 - p_c}{1 - p_c + p_c(1-s)} - \frac{nc}{nc + sp_c c} E(I_E^A(b_i)|c=0)$$

$$= \frac{nc}{nc + sp_c c} a - \frac{nc}{nc + sp_c c} E(I_E^A(b_i)|c=1) + \frac{1 - p_c}{1 - p_c + p_c(1-s)}$$

$$= \frac{1 - c}{1 - c + sp_c c} a - \frac{1 - c}{1 - c + sp_c c} E(I_E^A(b_i)|c=1) + \frac{1 - p_c}{1 - p_c + p_c(1-s)}. \tag{5}$$

Inverting the terms in equation (5), we can isolate a:

$$a = \frac{1 - c + sp_c c}{1 - c} E(\hat{a}) + E(I_E^A(b_i)|c=1)$$

$$- \frac{1 - p_c}{1 - p_c + p_c(1-s)} \times \frac{1 - c + sp_c c}{1 - c}$$

$$= \frac{1 - c + sp_c c}{1 - c} E(\hat{a}) - (1 - p_c) \frac{sp_c}{(1 - sp_c)(1 - c)}. \tag{6}$$

Equation (6) can be solved by explicitly deriving p_c and c: since we have $c(1 - sp_c) = \tilde{c}$, it follows that $c = \frac{\tilde{c}}{1 - sp_c}$. Specifically, from equations (3) and (4), we have

$$\Rightarrow p_c = \frac{1 - E(I_E^A(b_i)|\tilde{c}=1)}{1 - sE(I_E^A(b_i)|\tilde{c}=1)}.$$

Plugging the expressions for p_c and c in (6) yields an expression of a as a function of s:

$$a = \left(1 + \frac{s\tilde{c}(1 - E(I_E^A(b_i)|\tilde{c}=1))}{1 - s - \tilde{c}(1 - sE(I_E^A(b_i)|\tilde{c}=1))}\right) E(\hat{a})$$

$$- \frac{s(1-s)E(I_E^A(b_i)|\tilde{c}=1)(1 - E(I_E^A(b_i)|\tilde{c}=1))}{(1 - s - \tilde{c}(1 - sE(I_E^A(b_i)|\tilde{c}=1)))(1 - sE(I_E^A(b_i)|\tilde{c}=1))}.$$

First, when $s = 0$, $a = E(\hat{a})$. This means that when no co-ethnic defector passes as a non-coethnic, the OLS estimator of a is unbiased.

Second, for any strictly positive s, a is smaller than the OLS estimate of a. This means that, if there is some degree of co-ethnic passing, OLS systematically over-estimates the true effect of co-ethnic matching on $I_E^A(b_i)$.

Third, for any OLS estimate $\hat{a} > 0$, I can use this equation to back out the upper bound on the rate of passing that is consistent with the sign of the coefficient still be informative. Specifically, given an OLS estimate $\hat{a}$, the true value of a monotonically decreases in s. Thus, it follows that there exists $\bar{s}$ such that, for any $s > \bar{s}$, $a \leq 0$. That is, when the degree of passing s, is above that upper bound $\bar{s}$, selection bias arising from passing is so large

that a cannot be positive given the OLS estimate. How much passing can the true OLS estimate accommodate and still reflect a true effect of co-ethnic matching? The estimates from the data are $\hat{a} = 0.18$ (Table V, column (9)), $E(I_E^A(b_i)|\tilde{c} = 1) = 0.33$ (Table I, Panel 4, column $S\{C\}$) and $\tilde{c} = 0.41$ (Table I, columns $S\{C\}$ and $S\{N\}$, observations number) and since $c = \frac{\tilde{c}}{1 - sp_c} < 1$ must hold for s to be a fraction, we have: $\bar{s} = \frac{1 - \tilde{c}}{1 - \tilde{c}(E(I_E^A(b_i)|\tilde{c}_i = 1))} = 0.42$. That is, if $s < 0.42$, then we still have $a > 0$.

REFERENCES

OLEA, J., AND C. PFLUEGER (2013): "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, 31. [7]

STAIGER, D., AND J. H. STOCK (1997): "Instrumental Variables Regression With Weak Instruments," *Econometrica*, 65 (3), 557–586. [7]

Co-editor Fabrizio Zilibotti handled this manuscript.