BANCO CENTRAL DE RESERVA DEL PERÚ

Reducing Informality Using Two-Sided Incentives: Theory and Experiment

Francisco B. Galarza* y Fernando Requejo*

* Universidad del Pacífico

DT. N°. 2019-003
Serie de Documentos de Trabajo
Working Paper series
Febrero 2019

Reducing Informality Using Two-Sided Incentives: Theory and Experiment*

Francisco B. Galarza‡ & Fernando Requejo‡
Universidad del Pacífico
Lima, Peru

April 23, 2019

Abstract

We study the impact of two-sided incentives on the reduction of informality. We model those incentives using the notion of network externalities, which link the (formal or informal) merchant's profits to the type of customers they serve (formal or informal). Our theoretical framework yields two straightforward testable implications: the merchant will find more profitable to become formal (or informal), as long as more of their customers are formal (or informal); and, formal and informal commercial sectors may coexist in equilibrium. We test these hypotheses using data from a field experiment, conducted with micro and small enterprises in Lima, Peru. Our subjects had to choose, in a repeated fashion, among three 'platforms', which proxy for being formal, informal, or performing a reservation activity. We then changed the relative size of the network of formal vis-á-vis informal customers, in order to calculate the consumer's network externality. We find that the network externality is relatively large, a result that opens up the possibility to reduce commercial informality using two-sided incentives. Moreover, the platform choice between the formal and informal sectors is sensitive to risk preferences.

Key words: network externality, informality, two-sided incentives, experiments.

JEL Codes: C93, E26, O17.

* This paper is a revised version of the final report, entitled "Abordando la informalidad comercial usando incentivos bilaterales: un enfoque experimental", sponsored by the Consortium of Economic and Social Research (CIES, for its acronym in Spanish), and the National Superintendence of Customs and Tax Administration (SUNAT, for its acronym in Spanish) of Peru, under the research project PCT 1051. We thank Deyna Quiroz, Lucienne Neumann, Bruno Cáceres, Rodrigo Guerra, Tito Quispe and Willy Mak, for valuable assistance in the field. We are grateful to Alejandro Granda, two anonymous reviewers, and seminar participants at the Central Reserve Bank of Peru (BCRP) 36th Annual Meeting and the 2nd Wuhan Cherry Blossom Workshop in Experimental Economics (China), for useful comments to earlier versions of the paper. The views expressed herein are those of the authors only.
‡ Corresponding author. Department of Economics, Universidad del Pacífico. Postal Address: Universidad del Pacífico, Sánchez Cerro 2050. Jesús María, Lima, Peru. Phone: (51)1-219-0100. E-mail: galarza_fb@up.edu.pe.
‡ E-mail: fernandorequejoz@gmail.com.

1. Introduction

As in other Latin American countries, informality is widespread in Peru (De Soto, 1986; Schneider and Enste, 2000) and is regarded as a multidimensional problem, which affects multiple sectors of the economy, with different intensities (Loayza et al., 2010). Informality is particularly present in developing countries and tends to be a characteristic of the least productive agents, as reported by La Porta and Shleifer (2014). From the point of view of the Government, the main problems associated with the existence of an informal economy are the lack of regulation and tax evasion. In the case of Peruvian, Lahura (2016) estimates that the tax evasion of the financial system's informal customers amounts to 0.7% of the gross domestic product (GDP), as of 2014.

One of the economic sectors most affected by the informality is the commercial sector in big cities. The agglomeration of street vendors and other informal merchants is quite noticeable in large Peruvian cities, such as Lima (the capital), Trujillo, Chiclayo (Northern Peru), and Arequipa (Southern Peru), among many others. These traders have attained a level of coordination that enables them to successfully contest the traditional methods of Government's control and enforcement—inspections and seizure of merchandise (De Soto, 1986). Furthermore, this type of unregulated trade is directly associated with smuggling and reselling stolen goods, thus compounding other problems such as crime and scams (Mirus and Smith, 1997).

To date, there seems to be no consensus about the best way to eradicate the informality in the developing world. In the case of Peru, different types of policies have been proposed as a solution: (i) the development of programs of certification and improvements in access to micro-credit, under the influence of de Soto (1986);[1] (ii) the reduction in the cost of the paperwork; (iii) and, more recently, the reduction of the tax liabilities, with the premise that a decrease in taxes should reduce informality, but none of them has significantly reduced the size of the informal economy. Similar solutions have been proposed in other regions of the world, with comparable results (for the case of regulatory costs, see Bruhn & McKenzie, 2014).[2] Precisely, the complexity of the problem lies in that it is virtually impossible to define a proper functional form of all the variables that cause informality (Loayza, 2007). This, in turn, makes it difficult to propose a multidimensional effective policy.

It is important to remark that the existing literature has focused on supply-side interventions, leaving the demand-side approach understudied. In particular, to the best of our knowledge, no previous study has examined the effect of the network of people that interacts with an entrepreneur on her decision to become formal. In this paper, we address this problem from the perspective of a "network externality", which implies that decisions from the supply side (to become formal or informal) may be affected by the type, the intensity, and the number of links established with her suppliers and customers. In short, when the entrepreneurs have a high number of connections with individuals who are informal, those will have less incentive to become formal.

The concept of network externality also applies to consumers, obviously. In this paper, we do not model this side of the market, but for the theoretical framework we introduce later on, it is important to be aware of this connection. Given that consumers are influenced by a network externality coming from sellers and these ones see their decisions influenced by a network externality coming from consumers, these decisions fit in the context of two-sided incentives. This approach is based on the theory of two-sided markets and platform competition proposed by Caillaud and Jullien (2003), Rochet and Tirole (2003), and Armstrong (2006), among others. In those markets, a platform allows two groups to connect. These sides benefit from the interactions and such gains are larger the more they interact. Thus, the larger the number of individuals in one side, the bigger the benefit of the other side. This encourages the latter side to grow, which in turn, encourages

[1] Along the same line, Lavado and Campos (2017) highlight the lack of impact on informality of the reduction of the costs of formalization, at least for a large group of low-productivity entrepreneurs.

[2] Ulyssea (2018) finds that increasing regulation costs (and enforcement, in general) may reduce informality in Brazil. The effect of reducing entry costs, however, is negligible.

the former to grow. These dynamics continues until the marginal gain of the increase in the network size of both sides tends to zero.

As an example, consider markets such as credit cards and the apps to request a taxi. In the former case, if a larger number of merchants have a VISA, Mastercard or American Express POS, the cardholders of credit cards from those platforms will benefit as they can make more purchases and get other benefits from using their cards. As a result, a greater number of people will want to have these credit cards, which in turn encourages merchants to have the POS. Something similar happens with the taxi apps, the consoles of video games, crowdfunding, and so on.

We use the concept of network externality in the context of the decision to become a formal or informal merchant, in order to examine the role of two-sided incentives (given by the consumers' network externality) on the level of formalization. In particular, we provide the first measurement of such network externality for micro and small entrepreneurs (MSEs). Our hypothesis is that, using this type of incentives, the size of the formal market can increase. If this hypothesis is true, the government could increase the tax revenue without spending more in control, supervision, education or simplification of formalization procedures.

The rest of the paper is organized as follows. Section 2 introduces a theoretical model that guides our empirical section. Section 3 presents the experimental methodology we use to test our main hypotheses. Section 4 discusses our main experimental results and Section 5 concludes.

2. Model

As mentioned earlier, we propose a model using a two-sided market approach. Specifically, we exploit the modeling used in the journal and magazine markets,[3] which we have slightly modified, in order to account for the user's choice of formality or informality . We assume that there exist two types of users, consumers (indexed by "c") on the one-side, and merchants (indexed by "m"), on the other side. In the case of the consumers, we consider that each consumer buys a unique homogenous bundle of goods, which gives her a utility, s. For the merchants, we assume that they earn a utility, r, per bundle of goods sold. Moreover, any user, either a consumer or a merchant, must be "on board" on the formality or informality platforms and is restricted to interact with users on the other side that are "on board" in the same platform. That is to say that, formal (respectively informal) consumers interact with formal (respectively informal) merchants only.

Clearly, the merchant's utility depends on the number of interactions with their consumers. This means that a bigger share of formal consumers causes a bigger share of formal merchants (in a sense, a formal consumer may 'formalize' merchants) and the same is true for informal users. At the same time, since there is a finite number of consumers, each consuming a unique bundle of goods, the maximum consumer's market size "on board" on any platform is constant. As a consequence, the merchant's utility depends on their market share; which implies that a merchant's utility depends negatively on the number of merchants "on board" on the same platform (number of competitors). We use n_c^f, n_m^f, n_c^i and n_m^i to denote the number (size) of formal consumers, formal merchants, informal consumers and informal merchants (which we call networks), respectively. Hereafter, we will normalize the users' networks sizes to market shares.

Typically, in two-sided markets, two or more platforms compete in prices to "get both sides on board". In a similarly way, in our model, the formality and informality compete in "prices" in order to be the preferred transaction means. In other words, users must choose between the formality

[3] Since our aim in this paper is to highlight the importance of the consumer's choice of formality or informality on the merchant's formalization decision, the use of a general two-sided market approach results unnecessary. This model is a modified version of the one reported in our original paper (final report, sponsored by CIES and SUNAT).

platform and the informality platform. For simplicity, we could think of the formality price as taxes, while the informality price could be thought of as the expected punishment for being informal.[4] Both the formality and informality prices for merchants are defined as unit prices (per transaction, adjusted by their market share). For consumers, in turn, we consider that they pay a formality or informality price for their entire bundle of goods. We thus denote the formal prices charged to consumers and merchants as t_c and t_m, respectively, while the informal prices charged to informal consumers and merchants are denoted as d_c and d_m, respectively (we assume that those prices are fixed, for simplicity). Since these prices describe, in part, detection and punishment, it would be better to think of them as expected values.

Next, we will establish differentiation parameters between populations on each side. This will allow us to determine which users on each side will be part of the formality platform and the informality platform. In the case of the consumers, we assume that the formal platform has an additional cost, compared to the informal one. This cost could be understood as a transportation cost needed to reach formality, which varies among the consumer's population. It is convenient to think of this cost as a consumer's preference for formality, which makes it a parameter intimately related to her culture.[5] We thus define θ, as the distance a consumer is from formality, and l, as the per unit (of distance) cost, so that the total transportation cost will be, θl. For simplicity, we assume that θ is uniformly distributed in the unit interval $[0,1]$. See Figure 1a for the horizontal differentiation representation on the consumer side.

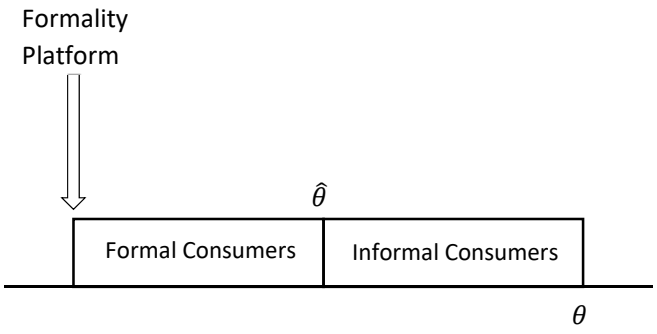


Figure 1a: Consumers

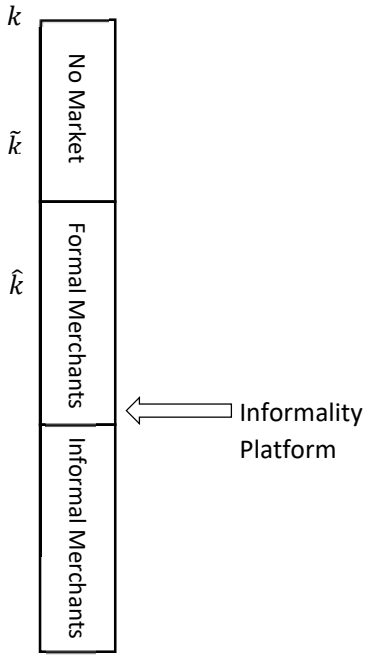


Figure 1b: Merchants

For the merchant side, we use a vertical differentiation approach. Unlike the consumer's side, our proposal for merchant side focuses on the merchant's risk preferences and not on cultural

[4] In general, in our approach, from the user's perspective, prices should be understood as any costs related to the interaction. In that sense, in the case of formality, it could include more than just taxes (e.g., the protection of the law, which reduces formality prices). The same applies to the informality, which has a lower level of protection of law (thus increasing the informality prices).

[5] We include it as a cost because, in general, is costly for a consumer to be formal, as it implies filing tax returns, waiting for receipts, and so on. Every consumer may value these costs differently.

aspects.[6] In this sense, our model considers that the most risk adverse merchants will be out of the market, a second group with lower risk aversion, will be formal, and the least risk adverse merchants will be informal. We thus denote k as the risk aversion parameter. For any merchant, a larger value of k implies higher risk aversion (see Figure 1b). For simplicity, again we assume that k is uniformly distributed in the unit interval $[0,1]$. We include the variance of the net returns (value r minus the platform price) as an indicator of the platform's risk. Then, we assume that the interaction between risk preferences and platform's risk is what negatively affects the merchant's utility (it is a simplification to assume that risk enters in logarithms in the merchants' utility function). We define the formality risk as σ_f^2 and the informality risk as σ_i^2. Putting all these pieces together, we have the following utility functions:

I. Formal consumers:

$$u_c^f = \begin{cases} s - t_c - \theta l, & n_m^f > 0 \\ 0, & n_m^f \leq 0 \end{cases} \dots (1)$$

II. Formal merchants:

$$u_m^f = \begin{cases} (r - t_m)\dfrac{n_c^f}{n_m^f} - kln(\sigma_f^2), & n_c^f > 0 \\ 0, & n_c^f \leq 0 \end{cases} \dots (2)$$

III. Informal consumers:

$$u_c^i = \begin{cases} s - d_c, & n_m^i > 0 \\ 0, & n_m^i \leq 0 \end{cases} \dots (3)$$

IV. Informal merchants:

$$u_m^i = \begin{cases} (r - d_m)\dfrac{n_c^i}{n_m^i} - kln(\sigma_i^2), & n_c^i > 0 \\ 0, & n_c^i \leq 0 \end{cases} \dots (4)$$

2.1 Networks

There exist three types of equilibria: pure formality, pure informality and the coexistence of formality and informality. The third case implies that one part of the market is formal and another is informal, at the same time, in both sides: $0 < n_c^f$, $0 < n_c^i$, $0 < n_m^f$ and $0 < n_m^i$. The pure cases can be relabeled as 'tipping equilibria' and there exist four scenarios in which this kind of equilibrium is attained: $n_c^f = 1$, $n_c^i = 1$, $n_m^i = 0$ and $n_m^f = 0$.

Network sizes are determined by user's decisions. Thus, any consumer j, with a preference parameter, θ_j, must decide between formality and informality, by comparing the utilities given by equations (1) and (3). If $u_c^f(\theta_j)$ is larger than $u_c^i(\theta_j)$, then being formal will be the best choice;

[6] An alternative specification would be to use a productivity argument, assuming that the most (least) productive merchants are forma (leave the market). Ulyssea (2018) follows this approach, though his analysis is broader and includes studying the firms' extensive margin (saving the registrations and entry costs) and intensive margin (hiring labor "off-the books") of informality. In this paper, we only examine the extensive margin.

otherwise, such consumer will choose informality. The characterization we propose here entails a complete market structure. This means that no consumer stays out of the market (see Figure 1a):

$$n_c^i + n_c^f = 1 \dots (5)$$

In order to find the formal consumer's share in the market, we make equations (1) and (3) equal, and find the value of $\hat{\theta}$, such that a consumer is indifferent between formality and informality. Then, since θ is uniformly distributed, we have:

$$\hat{\theta} = \frac{d_c - t_c}{l},$$

which means that the formality share is:

$$n_c^f = \frac{d_c - t_c}{l} \dots (6)$$

The informality share, n_c^i, can be inferred using equation (5). On the other hand, on the merchant's side, we follow the same procedure to find the formality and informality shares, n_m^f and n_m^i: find the indifferent merchant, $u_m^f(\hat{k}) = u_m^i(\hat{k})$, from equations (2) and (4). Then, since k is uniformly distributed, we have:

$$\hat{k} = (r - d_m)\frac{1 - \frac{d_c - t_c}{l}}{ln\left(\frac{\sigma_i^2}{\sigma_f^2}\right)n_m^i} - (r - t_m)\frac{\frac{d_c - t_c}{l}}{ln\left(\frac{\sigma_i^2}{\sigma_f^2}\right)n_m^f}$$

Using $\hat{k}$, and its uniform distribution, we find share of informal merchants, $n_m^i = \hat{k}$:

$$n_m^i = (r - d_m)\frac{1 - \frac{d_c - t_c}{l}}{ln\left(\frac{\sigma_i^2}{\sigma_f^2}\right)n_m^i} - (r - t_m)\frac{\frac{d_c - t_c}{l}}{ln\left(\frac{\sigma_i^2}{\sigma_f^2}\right)n_m^f} \dots (7)$$

As mentioned earlier, given that we assume that the least risk adverse merchants are informal, they 'live' in the interval $[0, \hat{k}]$. Moreover, any merchant for whom $k \geq \hat{k}$ will choose either to become formal (if she is not too risk averse) or to stay out of the market. Then, finding the formal merchants' share requires to determine the risk parameter that will make a merchant indifferent between formality and staying out of the market. If we assume that the latter choice has a utility of zero, we can find such threshold value, which we label as $\tilde{k}$ (see Figure 1b). Then, we have:

$$\tilde{k} = (r - t_m)\frac{n_c^f}{ln(\sigma_f^2)n_m^f}$$

Since k is uniformly distributed, we have that $n_m^f = \tilde{k} - \hat{k}$, whose full expression is as follows:

$$n_m^f = (r - t_m)\frac{\frac{d_c - t_c}{l}}{ln(\sigma_f^2)n_m^f} - (r - d_m)\frac{1 - \frac{d_c - t_c}{l}}{ln\left(\frac{\sigma_i^2}{\sigma_f^2}\right)n_m^i} + (r - t_m)\frac{\frac{d_c - t_c}{l}}{ln\left(\frac{\sigma_i^2}{\sigma_f^2}\right)n_m^f} \dots (8)$$

Using equations (7) and (8), we solve the system, and find the merchant's distribution between formality and informality in the market, n_m^f and n_m^i.

Now, in the tipping equilibria cases, if $n_c^f = 1$, there is no informal consumers ($n_c^i = 0$) and consequently, informality is not profitable for any merchant, because there are no consumers to offer products to. This happens only if $u_c^f > u_c^i$ is satisfied for every consumer in the market. Then, using equations (1) and (3) we have the following condition:

$$t_c \leq d_c - l \ldots (9)$$

At this point, knowing that u_c^i is zero (there are no informal consumers), comparing equation (2) with zero, we have:

$$n_m^f = \sqrt{\frac{(r - t_m)}{ln(\sigma_f^2)}} \ldots (10)$$

If $n_m^i = 0$, then consumers have no merchant to interact with on the informal merchants side, so that the consumers' utility of been informal is zero. Thus, we have two cases: maintain condition (5) or relax it. If condition (5) is satisfied, then $n_c^f = 1$ and on the merchant's side, the formality utility is compared with zero (the informal utility) only. As a result, we have that the merchant's side formal network follows the same value as the one indicated by equation (10).

On the other hand, if condition (5) is relaxed, then $n_c^f + n_c^i \leq 1$ and, considering that $n_c^i = 0$, we have $n_c^f \leq 1$. In this case, the consumer will compare her utility from formality with being out of the market. Then, we have:

$$n_c^f = \frac{s - t_c}{l} \ldots (11)$$

This characterization on the formal network on the consumers' side implies a modification on the results of the merchant's side. Because $u_m^i = 0$, the network on the informal merchant's side is zero, and the share of the formal sector is determined by comparing u_m^f with zero, but including equation (11). Then, we have:

$$n_m^f = \sqrt{\frac{(r - t_m)(s - t_c)}{ln(\sigma_f^2)l}} \ldots (12)$$

If $t_c > d_c$, then $n_c^i = 1$ and informality prevails on both sides. Also, the merchant's informal network is determined by using equation (4)—making it equal to zero. Then, we have:

$$n_m^i = \sqrt{\frac{(r - d_m)}{ln(\sigma_i^2)}} \ldots \ldots (13)$$

Finally, if $n_m^f = 0$, then $u_c^f = 0$. In consequence, equation (3) is compared to zero. Since there is no differentiation parameter, all consumers will be informal, so $n_c^i = 1$ and the merchant's informal network is characterized by equation (13).

2.2 Government's Income

For our purposes, we assume that the government maximizes tax revenues, choosing optimal formality prices on both sides (consumer and merchants). Income from the consumer's side is defined as the product of the number of formal consumers and the formality price, t_c. Similarly, the income from the merchant's side is defined as the product of the number of merchants, the merchant's share, and formality price, t_m. Then, the government's total income is equal to the sum of the consumer's side income plus the merchant's side income. We also consider a tax collection convex cost function. Then we have total tax revenue is defined as follows:

$$T = n_c^f t_c + \frac{n_c^f}{n_m^f} n_m^f t_m - C(t_c, t_m) \dots (14)$$

Simplifying the expression, we see that the merchant's network, n_m^f, does not affect tax revenues. Plugging equation (6) back into equation (14), we have:

$$T = \frac{d_c - t_c}{l}(t_c + t_m) - C(t_c, t_m),$$

which is the government's objective function. It is straightforward to show that T is a concave function and that its maximum value is unique. In this sense, the optimal tax decision could imply an optimal informality market level on both sides.[7]

3. Experiments

3.1 Experimental Design: Overview

A fundamental prediction of the model presented in section 2 is that, given the similarities between the competition that takes place between the formal and informal sectors, and the one between platforms in two-sided markets, one possible equilibrium outcome is that the informal sector coexists with the formal sector. As we saw earlier, depending on the parameters values, there exist three outcomes under a government's maximization of tax revenues: the market is completely informal ($n_m^i = 1$), the market completely formal ($n_m^f = 1$) and formality and informality coexist in the marketplace.[8] We designed a choice field experiment, in order to test which of those outcomes holds empirically for the MSEs, and whether, under a scenario in which the formal and informal sectors coexist, a two-sided incentive scheme, given by the consumers' network externality (which we will explain later in detail), could significantly reduce the share of the informal sector in the market. All this, for the commercial sector.

Our experiment replicates the decision a MSE may face while doing business, in terms of choosing the formal or the informal platform (or to stay out of the market, by choosing a reservation activity, instead), in a repeated fashion. We will thus observe how the change in the size of the consumers' network affects the relative size of the formal commercial sector (a variation of Figure 1a affecting Figure 1b). The repetition of the same "stage-game" during a finite (and known) number of times will capture the role of uncertainty in those investment decisions, and will make the MSEs understand (and feel) the costs and benefits of their choices over time. We thus assume that the decision of whether to become formal or informal is a choice (rather than an exit option), following McKenzie and Woodruff (2006).

3.2 Salient Features of the Entrepreneurial Behavior

[7] We ignore any welfare analysis here, as it is beyond the scope of this paper.
[8] A completely informal market is an unusual outcome, yet theoretically possible, if we consider an extremely high cost associated to tax collection.

A significant puzzle that remains unanswered in Development Economics has to do with the rationale behind the MSEs' choice to become a formal or informal merchant. What we do know is that reducing the entry costs may not suffice to reduce informality (Djankov et al., 2002).[9] In the case of Peru, Jaramillo (2013) offered a sample of MSEs the option to have the licensing cost (entry cost) completely paid in exchange for operating in the formal sector. Only one third of such sample took that option.[10] Why did so few merchants decide to become a formal merchant? Jaramillo conjectures that this ostensibly widespread reluctance to formalization may be because the MSEs foresee greater costs in the future (such as taxation, labor and product regulations), in comparison with staying in the informal sector. It thus seems reasonable to argue that risk aversion may play a role in the MSEs decisions:[11] if an informal MSE perceives the formalization as a risky choice, staying in the informal sector could be rational.

In order to include the notions of consumer's network externality, expected return and risk in the theoretical framework presented earlier, we need to make some adaptations. Let the "mark-up" that a merchant earns from each transaction be defined as:

$$E\left[\lambda^f\right] = r - t_m$$

$$E\left[\lambda^i\right] = r - d_m,$$

where the superscripts f and i denote the formal and informal sectors, respectively. Given the random nature of t_m and d_m, this mark-up will also be random. Thus, using equations (2) and (4), we can express the utilities of formal and informal merchants, respectively, as follows:

$$u_m^f = E\left[\lambda^f\right]\frac{n_c^f}{n_m^f} - kln\left(\sigma_f^2\right) \ldots (10)$$

$$u_m^i = E\left[\lambda^i\right]\frac{n_c^i}{n_m^i} - kln\left(\sigma_i^2\right) \ldots (11)$$

These functional forms are straightforward and easy to simulate in an experiment, since they only depend on the expected returns, the variance of those returns, and the user's network size ($n_{c,m}^{i,f}$). For simplicity, we assume that $\left(n_m^{i,f}\right)$ is fixed in each decision. Thus, for given expected returns and variances, the merchant's utilities will depend only on $\left(n_c^{i,f}\right)$, which will be our variable of interest (i.e., subject to experimental manipulation). The parameter k is unknown, and subject-specific.

We gathered survey information from a sample of 55 MSEs (merchants), in order to calculate those indicators. In particular, our risk (variance) indicator will be the probability to be scrutinized (visited by the tax authorities). That is, we only have an approximation of a binomial distribution of the parameter. Since we have information about the empirical expected returns and variances for the formal and informal merchants, we can construct a multinomial probability distribution, whose relative expected values and variances match the empirical values gathered from our survey. In this new distribution, we define α_1^f, α_2^f and α_3^f (α_1^i, α_2^i and α_3^i,), as the profitability that three particular realizations of the returns are obtained in the formal (informal) sector, denoted as λ_1^f,

[9] A study for Brazil, however, finds a positive effect of a large-scale formalization program that reduced taxes on formalization, once registration costs have been eliminated; although the net gains from such program are negative (Rocha et al., 2018).

[10] A related, experimental intervention was carried out in Brazil, where De Andrade et al. (2016) tested several strategies to encourage formalization, by reducing the costs of informality (via information campaigns, elimination of registration costs, and enforcement visits). Again, registration costs do not affect formalization, which is only affected by increasing the enforcement visits (costs to informality).

[11] Raunelli et al. (2016) find that a sample of Peruvian MSEs behave as risk averters.

λ_2^f, and λ_3^f (λ_1^i, λ_2^i, and λ_3^i). Thus, for the purpose of our experiments, the expected return for each transaction in each market will be given by:

$$\text{Formal: } E[\lambda^f] = \alpha_1^f \lambda_1^f + \alpha_2^f \lambda_2^f + \alpha_3^f \lambda_3^f$$
$$\text{Informal: } E[\lambda^i] = \alpha_1^i \lambda_1^i + \alpha_2^i \lambda_2^i + \alpha_3^i \lambda_3^i$$

And the variance of the returns in each market will be given by:

$$\text{Formal: } \sigma_f^2 = \alpha_1^f \left(\lambda_1^f - E[\lambda^f]\right)^2 + \alpha_2^f \left(\lambda_2^f - E[\lambda^f]\right)^2 + \alpha_3^f \left(\lambda_3^f - E[\lambda^f]\right)^2$$
$$\text{Informal: } \sigma_i^2 = \alpha_1^i \left(\lambda_1^i - E[\lambda^i]\right)^2 + \alpha_2^i \left(\lambda_2^i - E[\lambda^i]\right)^2 + \alpha_3^i \left(\lambda_3^i - E[\lambda^i]\right)^2$$

It is important to notice that, in our framework, the relative utilities from being a formal merchant with respect to being an informal merchant, depend on the relative returns and the relative variances (see equations (10) and (11)).

The next step is to choose several levels of consumer's network externality, n_c^f and n_c^i, in order to measure its effect on the size of the formal and informal commercial sectors, n_m^f and n_m^i. We arbitrarily chose five different network sizes with the intent to have a moderate level of variation. With these data, we constructed a payoff matrix that summarizes the choices at stake (see Table 1, where each column represents a particular network size). Note that, in addition to the formal and informal commercial sectors, we introduce a third option, which represents staying out of the market (or performing a reservation activity). For this alternative/platform, we arbitrarily assume that the returns are constant (zero variance), and lower than those in the formal or informal sectors (this is Figure 1b in practice).

Table 1
Experimental payoff matrices

Formal

	n_1^f	n_2^f	n_3^f	n_4^f	n_5^f
α_1^f	λ_1^f	λ_1^f	λ_1^f	λ_1^f	λ_1^f
α_2^f	λ_2^f	λ_2^f	λ_2^f	λ_2^f	λ_2^f
α_3^f	λ_3^f	λ_3^f	λ_3^f	λ_3^f	λ_3^f

Informal

	n_1^i	n_2^i	n_3^i	n_4^i	n_5^i
α_1^i	λ_1^i	λ_1^i	λ_1^i	λ_1^i	λ_1^i
α_2^i	λ_2^i	λ_2^i	λ_2^i	λ_2^i	λ_2^i
α_3^i	λ_3^i	λ_3^i	λ_3^i	λ_3^i	λ_3^i

Reservation activity

	n^n	n^n	n^n	n^n	n^n
α_1^n	λ^n	λ^n	λ^n	λ^n	λ^n
α_2^n	λ^n	λ^n	λ^n	λ^n	λ^n
α_3^n	λ^n	λ^n	λ^n	λ^n	λ^n

The major advantage of our approach is that we can capture the effect of the consumers' network externality in a very clean way during our experimental sessions, since we will only change the size of the consumers' market size. One last simplification we made has to do with the merchants' risk aversion, captured by the parameter k. We assume homogenous risk preferences across all MSEs belonging to a given platform, which is reflected by the fact that each return in the matrix is the same across different network sizes.

3.3 Constructing the Empirical Payoff Matrices

As mentioned earlier, we conducted a survey with a sample of 55 formal and informal MSEs, in 25 districts of Lima, the capital city of Peru. The design of the questions about expected returns was such that we can compute the relative expected returns. Thus, setting the expected return for an informal MSE to be 100, the return for a formal MSE would be 85.76,[12] with the relative return,

[12] This figure includes a deduction of 20%, which adjusts for a higher productivity that is usually attributed to the formal sector.

informal to formal, being 1.17. On the other hand, the variances of those returns were 78.66 (informal: σ_i^2) and 34.12 (formal: σ_f^2), with a relative variance, informal to formal, of 2.31.[13]

The payoffs used in our experiment replicate, as close as possible, those empirical values: expected returns: $E[\lambda^f] = 66.67$, $E[\lambda^i] = 76.67$, with $\frac{E[\lambda^i]}{E[\lambda^f]} = 1.15$; and variances: $\sigma_f^2 = 1{,}439$, $\sigma_i^2 = 3{,}489$, with relative variance, $\frac{\sigma^2}{\sigma^2} = 2.42$. Furthermore, we (arbitrarily) assume $E[\lambda^n] = 40$, $\sigma_n^2 = 0$. On the other hand, we chose five consumer's network sizes, which were presented in ascending order of the formal sector's share: size 1: $(n_1^f = 20,\ n_1^i = 80)$; size 2: $(n_2^f = 40,\ n_2^i = 60)$; size 3: $(n_3^f = 50, n_3^i = 50)$; size 4: $(n_4^f = 60, n_1^i = 40)$; size 5: $(n_5^f = 80, n_5^i = 20)$. The third choice always considers 50 consumers $(n^n = 50)$. In our design, then, as we go along with the experiment, as we increase (respectively, decrease) the relative size of the formal (informal) consumer's network, the formal (respectively, informal) sector becomes relatively more (less) profitable. Furthermore, in order to simplify the explanation of the returns during the experiment, we use the same proportion for each of the three realizations: i.e., $\alpha_t^j = \frac{1}{3}$, for all choice $j = i, f, n$ y each payoff realization, $t = 1{,}2{,}3$.

3.4 Experimental Procedures and Sample

We designed choice experiments, where the subjects (MSEs) choose, in a repeated fashion, one of the three platforms available. These choices were framed in the context of investing in buying a selling stand (*puesto de venta*) in one out of three shopping centers: *Polvos Amarillos* (representing the formal commercial sector); *Polvos Morados* (informal sector); and *Polvos Naranjas* (reservation activity). Those names were used, in order to introduce familiarity in the MSEs decisions. As mentioned earlier, we considered 5 network sizes (labeled as Stages A through E), and each one of them included 5 decision rounds.

Our sample of MSEs is mostly comprised of merchants working in one of the biggest shopping centers in Lima, *Polvos Azules*. The invitation to participate was sent via flyers. And we sent daily reminders using the speaker of the shopping center. We sampled 150, out of the roughly 2,000 selling stands that operate in *Polvos Azules*. Seven additional stands were sampled from retailers in downtown Lima. This sample is not random. Our experiments lasted for about 30 minutes, and cash winnings from participation ranged between 9PEN and 27PEN (about 3 to 8.5USD), with an average of 17PEN (see Table 3, last row). After the experiment, we had a debriefing section, in which we explained the main notions of risk and investment, applied to the decisions they made during the experiment. The field work was conducted during November of 2018.

The instructions (see Appendix 1A) were read aloud in each session and each participant received her worksheets (see Appendix 1B for a sample). We held many sessions in which the instructions were read 1 on 1. The figures in each payoff matrix were chosen to replicate the empirical data we gathered. In particular, a selling stand in *Polvos Amarillos* (the formal platform) has an expected return of 66.67 (average of 35, 45 and 120) and a variance of 1,439; and a stand in *Polvos Morados* (the informal platform) has an expected return of 76.67 (average of 30, 40 and 160) and a variance of 3,489. Also, the expected return of *Polvos Naranjas* (the reservation activity) is 40, with zero variance. Table 2 shows a sample payoff matrix, used for network size 1.

Furthermore, we simulated the realizations of the returns (the α_t^j in our model) using colored ping-pong balls, which our subjects randomly drew from a bag containing exactly three balls (represented by each row within each platform matrix). The red-colored ball represents having "a bad client" in a regular business day, the blue-colored denotes an "average client", while the green-

[13] We acknowledge that those figures might not be representative of the microentrepreneural sector. But, no such information seems to exist for Peru, or for any other country, for that matter.

colored one represents a "good client". These scenarios capture the uncertainty in their investment, which is out of their control.

Table 2
Sample Payoff Matrices used in the Experiment: Network size 1

Polvos Amarillos	Polvos Morados	Polvos Naranjas
n_1^f: 20 customers	n_1^i: 80 customers	n_1^n: 50 customers
RED: 35	RED: 30	RED: 40
BLUE: 45	BLUE: 40	BLUE: 40
GREEN: 120	GREEN: 160	GREEN: 40

For each network size, subjects played 5 consecutive rounds, after which a new network size was introduced. We presented five network sizes, for a total of 25 rounds. Each set of five rounds was intended to make the subjects experience the intertemporal effects from selecting their preferred investment choice. We also included a practice stage, with five rounds, which may clean up, at least partially, some learning effect from the initial periods of the actual experiment.

Our sample is composed of 157 subjects. Our typical subject is 34 years old; has more than 7 years of experience in the selling business; and is risk averse. Since we are particularly interested in analyzing whether platform choices are correlated with risk preferences, we included a hypothetical question (choosing 100PEN for sure or a 50-50 chance of getting 200PEN or nothing); our risk aversion indicator is a dummy variable for those choosing the safe choice ('risk averters'). Such proportion is 65%. Moreover, 47% of our subjects are males; 39% completed high school or less; 34% are married; 57% of merchants said they have a tax ID (which is our criterion to define a MSE as "formal") and 43% report to be the owner of the business.

Table 3
Basic descriptive statistics of our sample

	N	Average	Min	Max	Std. Dev.
Age (years)	157	33.71	18	62	10.53
Male	157	0.47	0	1	0.50
Risk averse (picked the safe choice)[1/]	157	0.65	0	1	0.48
Experience in business (years)	157	7.61	0.5	38	7.40
Completed high school or less	157	0.39	0	1	0.49
Married	157	0.34	0	1	0.48
Has a tax ID	157	0.57	0	1	0.49
Owns the business	157	0.43	0	1	0.49
Drew a "good client" (green ball)	157	0.33	0.12	0.60	0.09
Drew a "normal client" (blue ball)	157	0.34	0.12	0.64	0.10
Drew a "bad client" (red ball)	157	0.33	0.04	0.64	0.09
Chose Informal Platform (%)	157	0.49	0	1	0.18
Chose Formal Platform (%)	157	0.42	0	1	0.18
Chose Reservation Activity (%)	157	0.09	0	0.6	0.13
Winnings in experiment (PEN)	157	17.03	9	27	3.45

[1] Lottery consisted in choosing 100PEN with certainty, or a 50/50 chance of winning 200PEN or 0PEN. The stakes were hypothetical.

Table 3 also reports some experimental results, in terms of the type of clients 'drawn' by our subjects (one-third, on average) and the platform choice, which we will examine next. Furthermore, cash winnings from participating in the experiment are around 17PEN (USD5.3).

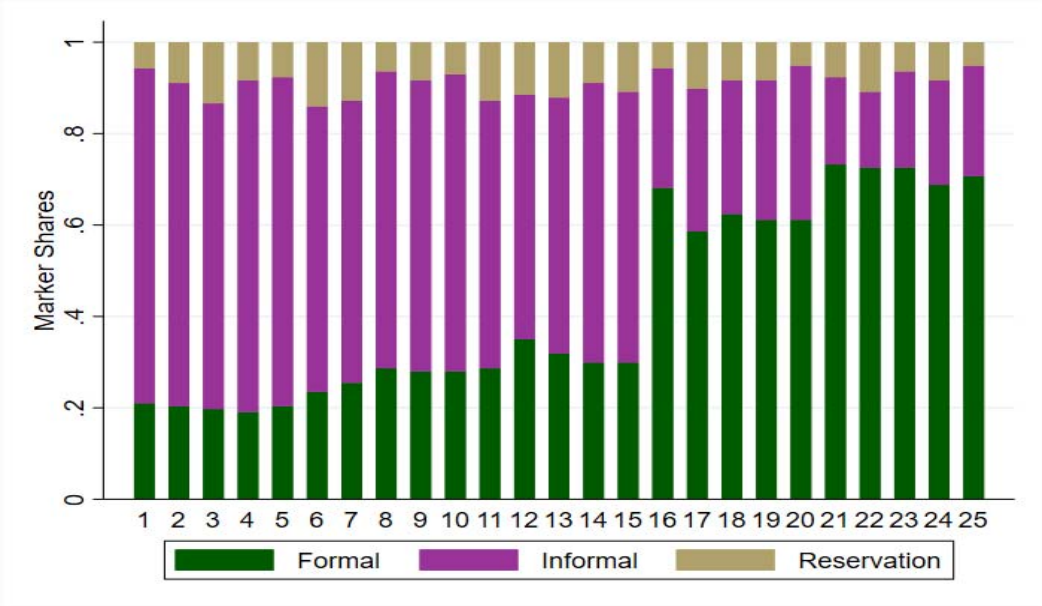
4. Experimental Results

We will next examine the consumer's network externality on the choice of the formal platform (vis-a-vis the informal platform) among our sample of MSEs.

4.1. Market Equilibrium and Network Externality

Our theoretical framework shows that the mixed equilibrium, where the formal and informal platforms coexist, is a possible equilibrium outcome. We find empirical support for this outcome. As can be inferred from Figure 1, in every round, the formal and informal sectors share around 91% of the market, on average. In particular, the formal platform shares 42% of the market, while the informal platform shares the remaining 49%, as shown in Table 3.

Figure 1
Merchants' Market Shares



As mentioned earlier, in order to capture the consumer's network externality, we changed such network size in the following way:

- Size 1 (rounds 1-5): 20, formal; 80, informal; 50, reservation activity
- Size 2 (rounds 6-10): 40, formal; 60, informal; 50, reservation activity
- Size 3 (rounds 11-15): 50, formal; 50, informal; 50, reservation activity
- Size 4 (rounds 16-20): 60, formal; 40, informal; 50, reservation activity
- Size 5 (rounds 21-25): 80, formal; 20, informal; 50, reservation activity

We will calculate the network externalities by computing the respective elasticities (using the share of the formal (or informal) consumer's network and the relative share (formal-to-informal or informal-to-formal shares).

Figure 2 plots the proportion of merchants who chose the formal (left panel) and informal (right panel) platforms in each of the 25 rounds. Remember that, from left to right, the number of formal clients (n^f) is increasing and that of informal clients (n^i) is decreasing, which makes the formal consumer's market relatively more profitable, on average. Figure 2 presents two clear results: (i) an overall positively- (negatively-) sloped formal (informal) merchants' market-share curve; and (ii) a marked shift in the level of the trends in round 16. As a whole, this figure shows that merchants react to the change in the consumers' network size, which is an effect that our experimental design aims to scrutinize.

Figure 2
Evolution of the Formal and Informal Merchants Network in Response to Changes in the Consumer's Network

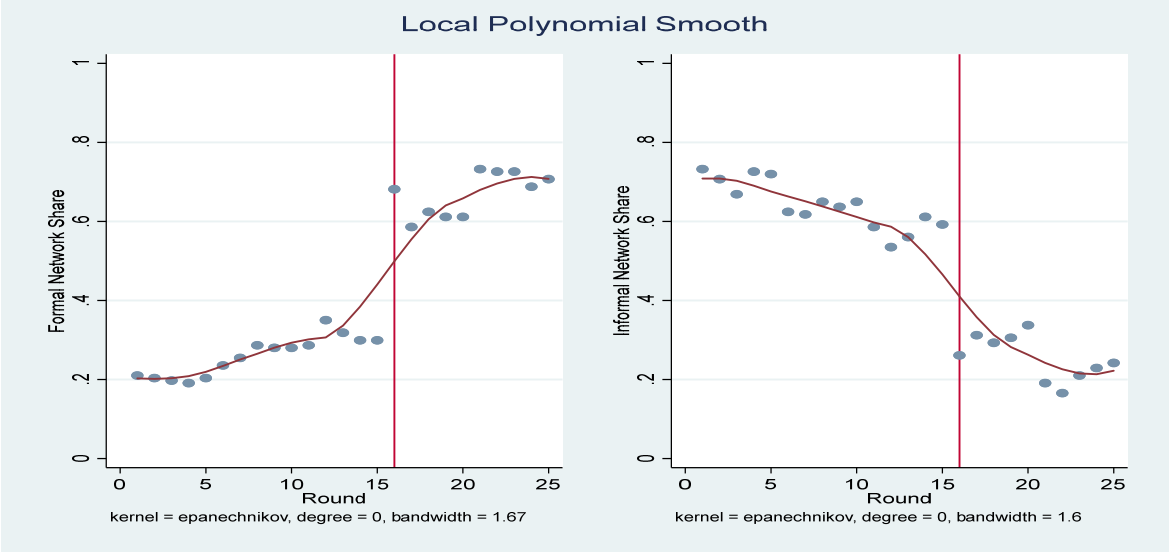
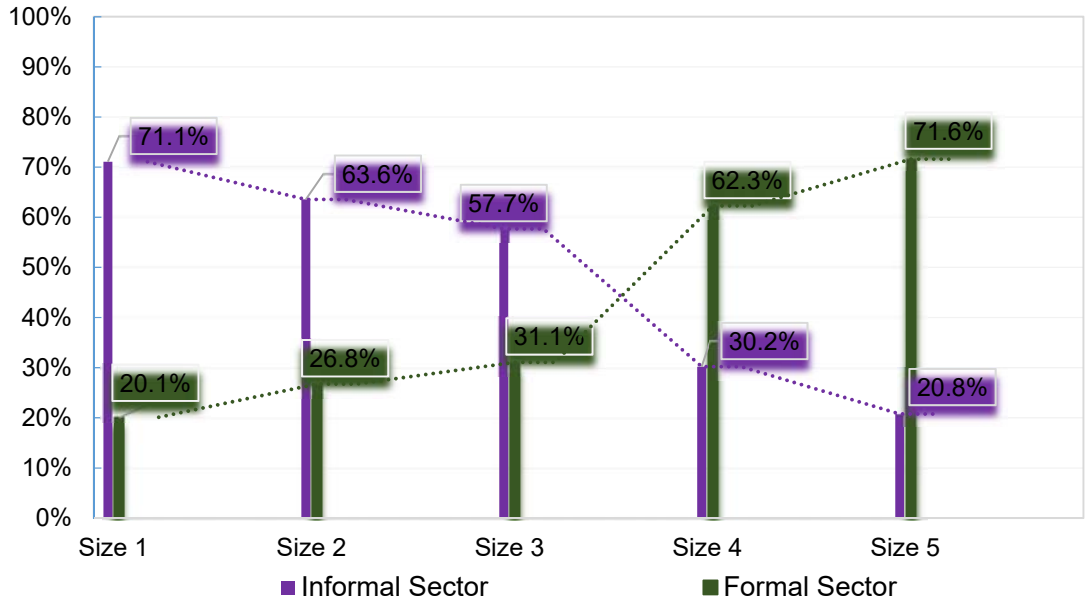


Figure 3 summarizes information from the previous two figures and displays the average share of the merchants' formal platform (increasing bars from left to right) and that of the merchants' informal platform (decreasing bars from the left) for each of the five consumer's network sizes. We will further explore both trends. Again, we see a marked change in trends, in network size 4 (rounds 16 to 20).

Figure 3
Evolution of the Merchant's Formal and Informal Network in Response to Changes in the Consumer's Network
(Average of each network size)



Because the relative size of the consumer's informal sector (with respect to the total) does not change uniformly across different consumer's network sizes, it is difficult to infer the magnitude of the consumer's network elasticity (how much the proportion of merchants who choose the formal or informal platform change when the share of the consumer's formal or informal sector grows by 1%), only by examining Figure 2.

As mentioned earlier, we calculate two types of consumer's network elasticities: using the absolute shares and the relative shares. First, we use the information of the percentage increase in the relative share of the formal (or informal) sector within each consumer's network size, as well as the proportion of merchants who choose the formal (informal) platform, shown in Figure 3. Table 4 displays the results. As seen below, the formal consumer's network elasticity of the formal merchant network average value is 1.611 (see column 3), while the corresponding figure for the informal consumer's network elasticity of the informal merchant network is smaller, but significantly larger than zero: 0.996 (column 6). Taken together, these results show an asymmetric externality of the formal and informal sectors, being the former a more elastic one (*increasing* the formal commercial sector seems easier than *reducing* the informal commercial sector). In either case, such externality starts to kick in in network size 4, when the formal consumers' network size is 60, the informal one is 40, and the reservation activity's is 50.

Table 4
Consumer's Network Elasticities of the Merchant's Network

Consumer's Network Size	Formal Merchant's Network (%)		Formal Consumer's Network (%)		***Formal* Network Elasticity (1) / (2)** (3)	Informal Merchant's Network (%)		Informal Consumer's Network (%)		***Informal* Network Elasticity (4) / (5)** (6)
	Share	Change (1)	Share	Change (2)		Share	Change (4)	Share	Change (5)	
Size 1	20.13		13.33			71.08		53.33		
Size 2	26.75	32.9	26.67	100.0	**0.329**	63.57	-10.6	40.00	-25.0	**0.423**
Size 3	31.08	16.2	33.33	25.0	**0.648**	57.71	-9.2	33.33	-16.7	**0.553**
Size 4	62.29	100.4	40.00	20.0	**5.020**	30.19	-47.7	26.67	-20.0	**2.384**
Size 5	71.59	14.9	53.33	33.3	**0.448**	20.76	-31.2	13.33	-50.0	**0.624**
Average					**1.611**					**0.996**

Second, we calculate the network elasticities using the percentage changes in the relative shares of formal-to-informal merchants and vice versa: we examine how much the relative share of the formal with respect to the informal merchant changes in response to a 1% change in the relative size of the formal with respect to the informal consumers; and vice versa. The results, shown in Table 5, confirm a relatively large formal-to-informal network elasticity (an average of 1.729, as seen in column 3) and a moderate informal-to-formal network elasticity (an average of 1.010, see column 6). Again, this elasticity is particularly large in network size 4 (greater than 5 and larger than 2, respectively).

Table 5
Consumer's Network Relative Shares Elasticities

Consumer's Network Size	Formal-to-informal				***Formal-to-Informal Network Elasticity*** **(1) / (2)** (3)	Informal-to-formal				***Informal-to-Formal Network*** **Elasticity (4) / (5)** (6)
	Merchants		Consumers			Merchants		Consumers		
	Relative Share	Change (1)	Relative Share	Change (2)		Relative Share	Change (4)	Relative Share	Change (5)	
Size 1	0.28		0.25			3.53		4.00		
Size 2	0.42	48.63	0.67	166.67	**0.292**	2.38	-32.72	1.50	-62.50	**0.523**
Size 3	0.54	27.99	1.00	50.00	**0.560**	1.86	-21.87	1.00	-33.33	**0.656**
Size 4	2.06	283.06	1.50	50.00	**5.661**	0.48	-73.89	0.67	-33.33	**2.217**
Size 5	3.45	67.10	4.00	166.67	**0.403**	0.29	-40.16	0.25	-62.50	**0.643**
Average					**1.729**					**1.010**

4.2 The Correlates of Platform Choice

In order to understand the rationale behind our subjects' choices, we use data about individual characteristics as potential correlates of their platform choices. In particular, given that subjects faced multiple choices, we run a multinomial logistic regression using the following empirical specification:

$$y_{it} = \alpha_0 + \alpha_1 Age_i + \alpha_2 Male_i + \alpha_3 Risk_i + \alpha_4 Experience_i + \alpha_5 Bad_{it} + \alpha_6 Normal_{it} + \alpha_7 X_i + \varepsilon_{it},$$

where y_{it} is the platform chosen by subject "i" at round "t": formal, informal, or reservation activity; Age_i represent the age in years; $Male_i$ is a dummy variable for males; $Risk_i$ is our measure of risk aversion; $Experience_i$ is the number of years in the selling business; Bad_{it} is a dummy for individual "i" drawing a red ping pong ball (a "bad client") in round t; $Normal_{it}$ is a dummy variable for individual "i" drawing a blue ping pong ball (a "normal, average client") in round t; and X_i is a set of individual-level control variables, including marital status, ownership of the business, and having a tax ID.

We run three specifications of the previous equation (results are reported in Table 6). None of the variables under scrutiny are significant for all three platforms. Thus, while age (experience) appears positively (negatively) correlated with choosing the reservation activity, risk averse MSEs tend to choose the formal platform rather than the informal platform. Furthermore, we wanted to examine the behavioral response of MSEs to the random draw of the type of clients (good ones, normal, or bad ones) they were serving. We find that the choice of the informal platform is negatively affected by drawing a "bad client" or a "normal client" (instead of drawing a "good client"). When this profit-reducing events happened, our subjects were more prone to choose the reservation, safer activity, a result that is consistent with the risk aversion behavior that our subjects seem to have. These results are robust to the inclusion of civil status, ownership of the business and having a tax ID (a minimum requirement to operate in the formal sector), as controls (see specification (3) below).

Table 6
Multinomial Logit Results: Platform Choice
(Marginal Effects)

	(1)			(2)			(3)		
	R.Activ.	Informal	Formal	R.Activ.	Informal	Formal	R.Activ.	Informal	Formal
Age (years)	0.002***	-0.001	-0.001	0.003***	-0.002	-0.001	0.003***	-0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Male	0.013	-0.006	-0.008	0.017	-0.010	-0.006	0.016	-0.012	-0.005
	(0.009)	(0.016)	(0.016)	(0.009)	(0.016)	(0.016)	(0.009)	(0.016)	(0.016)
Risk averse[1/]	-0.003	-0.036*	0.039*	-0.004	-0.035*	0.039*	0.001	-0.039*	0.037*
	(0.010)	(0.017)	(0.016)	(0.010)	(0.017)	(0.016)	(0.010)	(0.017)	(0.017)
Experience (years)	-0.002*	-0.000	0.002	-0.002*	0.000	0.002	-0.002*	-0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Drew a "bad client" (red ball)[2/]	0.008	-0.039*	0.031	0.008	-0.039*	0.031	0.008	-0.040*	0.032
	(0.012)	(0.020)	(0.019)	(0.012)	(0.020)	(0.019)	(0.012)	(0.020)	(0.019)
Drew a "normal client" (blue ball)[2/]	0.044***	-0.057**	0.014	0.044***	-0.058**	0.014	0.044***	-0.059**	0.015
	(0.011)	(0.019)	(0.019)	(0.011)	(0.019)	(0.019)	(0.011)	(0.019)	(0.019)
Observations	3925	3925	3925	3925	3925	3925	3925	3925	3925
Controls		No			Married			Married, Ownership, tax ID	

[1/] This is a dummy that takes the value of 1, if the MSE chose a hypothetical 100PEN for certain, instead of a 50/50 chance to earn 200PEN or 0PEN.
[2/] The omitted category is drawing a “good client” (green ball).
Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

We perform two robustness checks. First, we include an education dummy variable for those completing secondary education or lower (see Appendix Table 2A). Second, we include dummy variables for each round starting a new network size (which captures the network externality we described earlier) and also district of residence fixed effects, which may capture the MSEs socio-economic status (Appendix Table 2B). In general, the main results remain unaltered in both scenarios.

5. Conclusion

At first glance, an analysis of the consumer’s network externality may seem trivial, since it is sensible that any MSE will prefer to have a greater number of clients under a scheme in which they all give the same marginal profit. In this paper, we replicated two fundamental empirical features related to the MSEs’ expected returns and risk.

Our main finding is that the consumer’s network externality is significant; a result that suggests that there is an important room for the use of two-sided incentives in the reduction of the informal commercial sector in Peru. This country, like many other emerging economies, seem to be trapped in a vicious circle, in which the network of informal merchants has a considerable size. Our findings suggest that a plausible way out of such vicious circle may be to provide incentives to encourage a minimum number of consumers to be formal. Using our figures, such minimum size of network of formal consumers is around 60%. Unfortunately, we do not know the actual size of the formal consumers’ network, but it might be far below from 50% (we are speculating here, obviously).

While our paper suffers from all shortcomings attributable to experimental studies (such as the lack of external validity), we do believe that our results open up a potentially fruitful research on informal economies. It is part of our research agenda to test the effect of particular two-sided

incentives on the formalization levels, and to examine the sensitivity of platform choices to the relative returns and variances.

References

Allingham, M. and A. Sandmo (1972). “Income Tax Evasion: A Theoretical Analysis,” *Journal of Public Economics*, 1(3-4): 323-338.

Bruhn, M. and D. McKenzie (2014). “Entry regulation and the formalization of microenterprises in developing countries,” *The World Bank Research Observer*, 29 (2): 186–201.

De Andrade, G. H., M. Bruhn, & D. McKenzie (2016). “A Helping Hand or the Long Arm of the Law? Experimental Evidence on What Governments Can Do to Formalize Firms,” *The World Bank Economic Review*. 30(1): 24-54.

De Soto, H. (1986). *The Other Path: The Economic Answer to Terrorism.* Lima: Barranco.

Djankov, S., R. L. Porta, F. Lopez-De-Silanes, and A. Shleifer (2002). “The regulation of entry,” *The Quarterly Journal of Economics*, 117(1): 1–37.

Jaramillo, M**.** (2013). “Is There Demand for Formality among Informal Firms? Evidence from Microfirms in Downtown Lima.” *Avances de Investigación 13*. Lima: GRADE.

Lahura, E. (2016). “Sistema financiero, informalidad y evasión tributaria en el Perú,” *Documentos de Trabajo* No. 2016-11. Lima: Banco Central de Reserva del Perú.

La Porta, R. and A. Shleifer (2014). "Informality and Development," *Journal of Economic Perspectives*, 28(3): 109-26.

Lavado, P. and D. Campos (2017). “Empleo e informalidad.” En: Anderson, J. et al. (2017). *Balance de investigación en políticas públicas 2011-2016 y agenda de investigación 2017-2021*. Lima: Consorcio de Investigación Económica y Social (CIES).

Loayza, N. (1996). “The Economics of the Informal Sector: A Simple Model and Some Evidence from Latin America,” *Carnegie-Rochester Conference Series on Public Policy*, 45(1): 129-162.

Loayza, N. (2007). “The Causes and Consequences of Informality in Peru,” *Banco Central de Reserva Del Perú DT. N°* 2007-018. Lima: Banco Central de Reserva del Perú.

Loayza, N., L. Serven and N. Sugawara (2010). “Informality in Latin America and the Caribbean”, *Policy Research Working Paper Series* 4888. Washington, DC: Banco Mundial.

Mirus, R. and R. Smith (1997). “Canada’s Underground Economy: Measurement and Implications”. In: O. Lippert and M. Walker (eds.), *The Underground Economy: Global Evidence of its Size and Impact*. Vancouver, Canada: The Fraser Institute.

McKenzie D. and C. Woodruff (2006). “Do Entry Costs Provide an Empirical Basis for Poverty Traps?,” *Economic Development and Cultural Change*, 55(1): 3-42.

Raunelli, J. C., M. Power and F. Galarza (2016). “Why do entrepreneurs leave the market? An explanation from experimental economics in Peru”. *Revista Economía*, 39(77): 187-228. Lima: PUCP.

Rocha, R., G. Ulyssea and L. Rachter (2018). “Do lower taxes reduce informality? Evidence from Brazil,” Jo*urnal of Development Economics*, 134: 28-49.

Rochet, J.-C. and J. Tirole (2003). “Platform Competition in Two-Sided Markets.” *Journal of the European Economic Association*, 1(4): 990–1029.

Schneider, F. and D.H. Enste (2000). “Shadow Economies: Size, Causes, and Consequences,” *Journal of Economic Literature*, 38(1): 77-114.

Ulyssea, G. (2018). “Firms, Informality, and Development: Theory and Evidence from Brazil,” *American Economic Review*, 108(8): 2015-2047.

Appendix 1: Experimental Materials

Appendix 1A
Instructions for the Experiments

Good morning / good afternoon everyone. Thank you very much for attending this session and thanks to managers of the *Polvos Azules* Shopping Center, for allowing us to be here. As we announced by the shopping center's speaker and the flyers delivered in your posts, today's session has a training component, which we will develop through a dynamic that will last for about HALF AN HOUR. The activity will consist of choices that you must make. These choices will allow you to earn points. Those points will be then exchanged by cash at the end of the session. The more points you earn; the more money you will receive.

This session is part of a study with merchants in Lima that we are conducting at the Universidad del Pacífico. The objective of the study is to learn you, merchants, make decisions under circumstances in which the outcomes are uncertain. As you will know, little is still known about the commercial sector. The aim of our study is to learn a bit more about this important business sector.

Before we start with the activity that will allow you to earn points, we would like to practice, to make sure that the instructions are well understood. In the next activity, you will not earn points, but you will learn what to do in the rest of the session, in order to earn points.

Let's start. Think that you want to open a store in a shopping center and must choose among three shopping centers that are located one next to another, on the same block of a main avenue of the city. We will call them: *Polvos Amarillos*, *Polvos Morados* and *Polvos Naranjas*. As usual, your objective when you open a store is to earn as much money as possible. Throughout today's session, when making your decisions, please consider only the information we give you.

Practice Stage

If you choose to open your store in *POLVOS AMARILLOS*, you will have 50 customers each working day. Think of this number of clients as an average number of buyers that a store in *Polvos Amarillos* has in a typical day of the month. As you know, the net profit (revenue minus cost) for EACH client, depends on several factors, including the item that is purchased (which can be cheap or expensive) and how many items each client purchases (which depends on how much money the customer brings to buy). To all of these factors we will refer to as "Luck". We will consider three types of luck: good luck (represented by the color green); normal luck (blue), and bad luck (red). Then, if you decide to open your store in *Polvos Amarillos*, you will earn 120 points for EACH client, if you have 'good' luck; 80 points for EACH client, if you have 'normal' luck; and 30 points for EACH client, if you have 'bad' luck. Thus, your total net profits will be 120 x 50 = 6,000 points, if you have good luck; 80 x 50 = 4,000 points, if you have normal luck; and 30 x 50 = 1,500 points, if you have bad luck.

Do you have any questions so far? ... [Wait]

On the other hand, if you choose to open your store in POLVOS MORADOS, each working day you will also have 50 customers. And, your earnings will be as follows: 100 points for EACH customer, if you have good luck; 80 points for EACH client, if you have normal luck; and 50 points for each customer, if you have bad luck. Thus, your total net profits will be... 100 x 50 = 5,000 points, if you have good luck; 80 x 50 = 4,000 points, if you have normal luck; and 50 x 50 = 2,500 points, if you have bad luck.

Do you have any questions?... [Wait]

Finally, if you choose to open your store in POLVOS NARANJAS, each working day you will also 50 customers and will always earn 45 points for EACH customer. Thus, your TOTAL net profits will always be 45 x 50 = 2,250 points. Note that this is the only case in which all types of luck give you the same net profit for EACH customer. The earnings per customer, in this case, are "safe".

Do you have any questions?... [Wait]

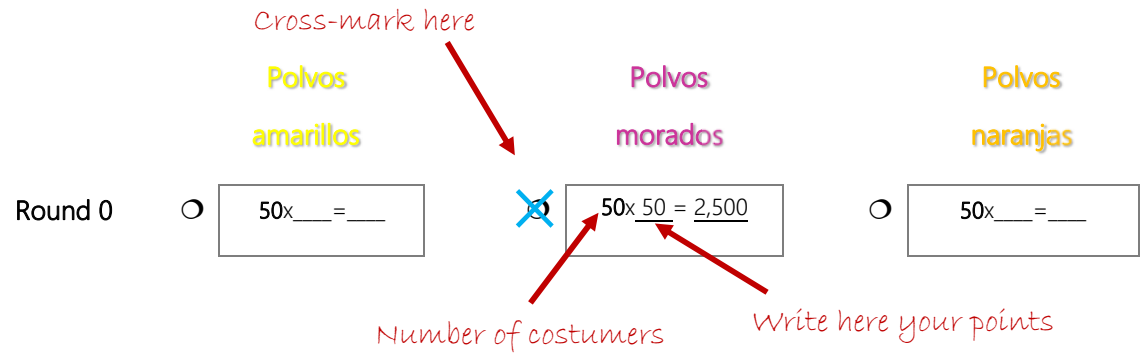
Earnings for each client from each shopping center are shown in the table below.

Polvos Amarillos	Polvos Morados	Polvos Naranjas
50 customers	50 customers	50 customers
RED : 30	RED : 50	RED : 45
BLUE : 80	BLUE : 80	BLUE : 45
GREEN : 120	GREEN : 100	GREEN : 45

What is this practice activity about? You have to choose at which shopping center you would like to open a store, only considering the information we just gave you. As we said earlier, once you have chosen if you want to open your store in *Polvos Amarillos*, *Polvos Morados* or *Polvos Naranjas*, your earnings for each client will depend on luck and can take one of three values in each case. In order to simulate “luck”, you will draw a colored ball from a bag containing exactly 1 green ball (good luck), 1 blue ball (normal luck) and 1 red ball (bad luck).
[Assistants: show the bag and the three colored balls ...]

Below is an example of what to do if someone decides to open a store in *Polvos Amarillos* and draws a normal luck (blue-colored ball).

Example



Do you have any questions about the instructions and earnings so far? [Wait...]

Okay, let's start the practice rounds. Remember that you must choose whether to open your store in *Polvos Amarillos*, *Polvos Morados* o *Polvos Naranjas* in each of the following 5 rounds. As we said earlier, at this stage, those 3 shopping malls have 50 customers; and the net profits from each Mall are the ones explained above.

Use your worksheet #1 (see table below), to cross mark the shopping center of your choosing in round 1. Then, wait to draw a colored ball, which will tell you the net profit for EACH customer. With this information known, please multiply the number of customers you have by your net profit per customer in each round. Once finished, move to the next round. Remember that you can only pick your shopping center on the next round if you already drew the ball from the previous round.

	Polvos Amarillos	Polvos Morados	Polvos Naranjas
Round 1	❍ 50 x___ = _____	❍ 50 x___ = _____	❍ 50 x___ = _____

Before moving on to round 2, draw a colored ball from the bag.

Round 2	❍ 50 x___ = _____	❍ 50 x___ = _____	❍ 50 x___ = _____

Before moving on to round 3, draw a colored ball from the bag.

Round 3	❍ 50 x___ = _____	❍ 50 x___ = _____	❍ 50 x___ = _____

Before moving on to round 4, draw a colored ball from the bag.

Round 4	❍ 50 x___ = _____	❍ 50 x___ = _____	❍ 50 x___ = _____

Before moving on to round 5, draw a colored ball from the bag.

Round 5	❍ 50 x___ = _____	❍ 50 x___ = _____	❍ 50 x___ = _____

Total points, Practice Stage	

Please add the points earned in the previous table. Later on, when we carry out the activity for cash, we will convert these points into cash using the following exchange rate: for every 6,000 points, we'll give you 1PEN in cash, paid at the end of the session.

But, once again, this is just a Practice Stage, aimed at making sure we all understand the instructions.

[…]

We will now start the stages where you will earn money. There hay 5 stages (A through E), with 5 rounds each. Let's begin stage A, which has 5 rounds. The decisions that you will make will always be the same: to choose in which shopping center to open your next store: in *Polvos Amarillos*,

Polvos Morados or *Polvos Naranjas*, taking into account the number of customers and the net profits that yield EACH customer.

Stage A

Please turn to your worksheet #2.

At this Stage, if you choose to open your store in *Polvos Amarillos*, in each working day, you will have 20 customers. After your choice of shopping center, you will draw a colored-ball from a bag containing 3 balls: 1 red, 1 blue and 1 green. If you draw the red ball, you'll earn 35 points for each customer that you have; if you draw the blue ball, you will earn 45 points for each customer; and if you draw the green ball, you will earn 120 points for every customer. Thus, your TOTAL earnings will be: 35*20 (= 700 points), in the first case; 45*20 (= 900 points), in the second case; and 120*20 (= 2400 points), in the third case.

Now, if you choose to open your store in *Polvos Morados,* in each working day, you will have 80 customers. After your choice of shopping center, you will draw a colored-ball from a bag containing 3 balls: 1 red, 1 blue and 1 green. If you draw the red ball, you'll earn 30 points for each customer that you have; if you draw the blue ball, you will earn 40 points for each customer; and if you draw the green ball, you will earn 160 points for every customer. Thus, your TOTAL earnings will be: 30*80 (= 2400 points), in the first case; 40*80 (= 3200 points), in the second case; and 160*80 (= 12800 points), in the third case.

Finally, if you choose to open your store in *Polvos Naranjas*, in each working day, you will have 50 customers and will earn 40 points for each client (for all types of luck). Thus, your TOTAL earnings will always be 50*40 = 2000 points.

As you can see, choosing a shopping center affects both the number of clients you will have and the net profits for each customer.

Payoff Matrix

Polvos Amarillos	Polvos morados	Polvos Naranjas
20 customers	80 customers	50 customers
RED : 35	RED : 30	RED : 40
BLUE : 45	BLUE : 40	BLUE : 40
GREEN : 120	GREEN : 160	GREEN : 40

Once you have chosen if you want to open your store in *Polvos Amarillos*, *Polvos Morados* or *Polvos Naranjas*, we will determine your earnings at random. For example, if you choose *Polvos Amarillos* and your randomly selected per-customer-net profit turns out to be 45 (because you drew the blue ball), your TOTAL net profit will be 45*20 = 900 points. On the other hand, if you had chosen *Polvos Morados* and your randomly selected per-customer-net profit turned out to be 40 (you drew the blue ball), your TOTAL net profit would be 40*80 = 3200 points. Finally, if you choose to open a store in *Polvos Naranjas*, your per-customer-net profit will always be 40 points, and your TOTAL net profit will always be 40*50 = 2000 points.

Do you have any questions about the instructions or earnings so far? [Wait...]

Okay, now let's start the activity. Remember that you must choose whether to open your store in *Polvos Amarillos*, *Polvos Morados* or *Polvos Naranjas*, and that your earnings from each election are those mentioned earlier. In your worksheet #2, cross-mark your selected shopping center in round 1. Once you draw a colored ball, write your per-customer net profit and your total net profit. Then, move on to the next round. Remember that you can only pick your shopping center in the next round if you already drew a colored ball from the previous round.

	Polvos Amarillos	Polvos morados	Polvos Naranjas
Round 1	❍ 20 x___ = _____	❍ 80 x___ = _____	❍ 50 x___ = _____

Before moving on to round 2, draw a colored ball from the bag.

Round 2	❍ 20 x___ = _____	❍ 80 x___ = _____	❍ 50 x___ = _____

Before moving on to round 3, draw a colored ball from the bag.

Round 3	❍ 20 x___ = _____	❍ 80 x___ = _____	❍ 50 x___ = _____

Before moving on to round 4, draw a colored ball from the bag.

Round 4	❍ 20 x___ = _____	❍ 80 x___ = _____	❍ 50 x___ = _____

Before moving on to round 5, draw a colored ball from the bag.

Round 5	❍ 20 x___ = _____	❍ 80 x___ = _____	❍ 50 x___ = _____

Total Stage A	

Please, add all the total net profits earned in each round. Remember that, for every 6,000 points, you are earning 1PEN in cash, which will be paid at the end of the experiment.
[…]

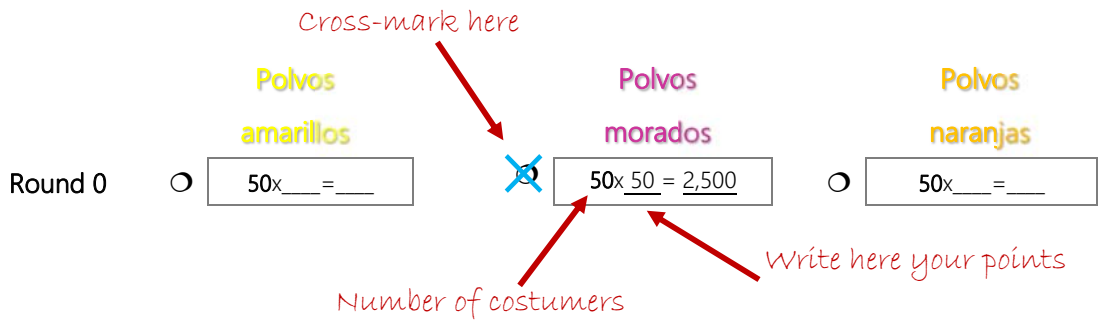
Similar instructions were provided for Stages B through E.

Appendix 1B
Worksheet Samples for Practice Rounds and Stage A—Network Size 1

Payoff Matrix: Practice Rounds

Polvos amarillos	Polvos morados	Polvos naranjas
50 customers	50 customers	50 customers
RED: 30	RED: 50	RED: 45
BLUE: 80	BLUE: 80	BLUE: 45
GREEN: 120	GREEN: 100	GREEN: 45

Example



Practice Rounds

	Polvos amarillos	Polvos morados	Polvos naranjas
Round 1	❍ 50x____=_____	❍ 50x____=_____	❍ 50x____=_____

Before moving on to round 2, draw a colored ball from the bag.

Round 2	❍ 50x____=_____	❍ 50x____=_____	❍ 50x____=_____

Before moving on to round 3, draw a colored ball from the bag.

Round 3	❍ 50x____=_____	❍ 50x____=_____	❍ 50x____=_____

Before moving on to round 4, draw a colored ball from the bag.

Round 4	❍ 50x____=_____	❍ 50x____=_____	❍ 50x____=_____

Before moving on to round 5, draw a colored ball from the bag.

Round 5	❍ 50x____=_____	❍ 50x____=_____	❍ 50x____=_____

Total Points Practice Stage	

Payoff Matrix: Stage A—Network Size 1

Polvos amarillos	Polvos morados	Polvos naranjas
20 customers	80 customers	50 customers
RED: 35	RED: 30	RED: 40
BLUE: 45	BLUE: 40	BLUE: 40
GREEN: 120	GREEN: 160	GREEN: 40

Stage A

	Polvos amarillos	Polvos morados	Polvos naranjas
Round 1	❍ 20x____=_____	❍ 80x____=_____	❍ 50x____=_____

Before moving on to round 2, draw a colored ball from the bag.

Round 2	❍ 20x____=_____	❍ 80x____=_____	❍ 50x____=_____

Before moving on to round 3, draw a colored ball from the bag.

Round 3	❍ 20x____=_____	❍ 80x____=_____	❍ 50x____=_____

Before moving on to round 4, draw a colored ball from the bag.

Round 4	❍ 20x____=_____	❍ 80x____=_____	❍ 50x____=_____

Before moving on to round 5, draw a colored ball from the bag.

Round 5	❍ 20x____=_____	❍ 80x____=_____	❍ 50x____=_____

Total Stage A	

Remember that you will earn 1PEN in cash for every 6,000 points.

Appendix Table 2A:

Multinomial Logit Results: Platform Choice
(Marginal Effects)

	(1)			(2)			(3)		
	R. Activity	Informal	Formal	R. Activity	Informal	Formal	R. Activity	Informal	Formal
Age (years)	0.002***	-0.002	-0.001	0.003***	-0.002*	-0.000	0.003***	-0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Male	0.015	-0.009	-0.006	0.019	-0.015	-0.004	0.018	-0.017	-0.002
	(0.010)	(0.016)	(0.016)	(0.010)	(0.016)	(0.016)	(0.010)	(0.016)	(0.016)
Risk aversion[1/]	-0.004	-0.035*	0.039*	-0.005	-0.034*	0.039*	0.000	-0.037*	0.036*
	(0.009)	(0.017)	(0.016)	(0.010)	(0.017)	(0.016)	(0.010)	(0.017)	(0.017)
Experience (years)	-0.002*	0.000	0.002	-0.003**	0.001	0.002	-0.002*	0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
High School or Lower	0.010	-0.030	0.020	0.013	-0.037*	0.023	0.013	-0.039*	0.027
	(0.010)	(0.017)	(0.017)	(0.010)	(0.017)	(0.017)	(0.010)	(0.017)	(0.017)
Drew a "bad client" (red ball)[2/]	0.008	-0.038*	0.030	0.008	-0.038	0.030	0.007	-0.038*	0.031
	(0.012)	(0.020)	(0.019)	(0.012)	(0.020)	(0.019)	(0.012)	(0.019)	(0.019)
Drew a "normal client" (blue ball)[2/]	0.044***	-0.057**	0.013	0.044***	-0.058**	0.014	0.044***	-0.059**	0.015
	(0.011)	(0.019)	(0.019)	(0.011)	(0.019)	(0.019)	(0.011)	(0.019)	(0.019)
Observations	3925	3925	3925	3925	3925	3925	3925	3925	3925
Controls		N			Married			Married, Owner, tax ID	

[1/] This is a dummy that takes the value of 1, if the MSE chose a hypothetical 100PEN for certain, instead of a 50/50 chance to earn 200PEN or 0PEN.
[2/] The omitted category is drawing a “good client” (green ball).
Robust standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix Table 2B:

Multinomial Logit Results: Platform Choice

(Marginal Effects)

	(1)			(2)			(3)			(4)		
	R.Activ.	Informal	Formal	R.Activ.	Informal	Formal	R.Activ.	Informal	Formal	R.Activ.	Informal	Formal
Age (years)	0.002***	-0.002	-0.001	0.003***	-0.002*	-0.001	0.003***	-0.002	-0.001	0.003***	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Male	0.014	-0.009	-0.006	0.018	-0.015	-0.004	0.018	-0.016	-0.002	0.020*	-0.023	0.002
	(0.010)	(0.016)	(0.016)	(0.010)	(0.016)	(0.016)	(0.010)	(0.016)	(0.016)	(0.010)	(0.017)	(0.016)
Risk Aversion[1/]	-0.004	-0.035*	0.039*	-0.004	-0.034*	0.038*	0.001	-0.037*	0.036*	0.006	-0.040*	0.034*
	(0.009)	(0.017)	(0.016)	(0.009)	(0.016)	(0.016)	(0.010)	(0.017)	(0.016)	(0.010)	(0.017)	(0.017)
Experience (years)	-0.002*	0.000	0.002	-0.003**	0.001	0.002	-0.002*	0.000	0.002	-0.002*	0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
High School or lower	0.010	-0.030	0.020	0.013	-0.036*	0.023	0.013	-0.039*	0.027	0.015	-0.040*	0.026
	(0.010)	(0.017)	(0.017)	(0.010)	(0.017)	(0.017)	(0.010)	(0.017)	(0.017)	(0.010)	(0.017)	(0.017)
Drew a "bad client" (red ball)[2/]	0.008	-0.040*	0.031	0.008	-0.039*	0.031	0.007	-0.040*	0.032	0.007	-0.040*	0.033
	(0.012)	(0.019)	(0.019)	(0.012)	(0.019)	(0.019)	(0.012)	(0.019)	(0.019)	(0.012)	(0.019)	(0.019)
Drew a "normal client" (blue ball)[2/]	0.044***	-0.054**	0.010	0.044***	-0.055**	0.010	0.044***	-0.056**	0.012	0.044***	-0.055**	0.010
	(0.011)	(0.019)	(0.019)	(0.011)	(0.019)	(0.019)	(0.011)	(0.019)	(0.019)	(0.011)	(0.019)	(0.019)
Round 6	0.049*	0.144***	-0.193***	0.049*	0.144***	-0.193***	0.049*	0.144***	-0.192***	0.048*	0.144***	-0.192***
	(0.020)	(0.043)	(0.045)	(0.020)	(0.042)	(0.044)	(0.020)	(0.042)	(0.045)	(0.020)	(0.042)	(0.044)
Round 11	0.036	0.094*	-0.130**	0.036	0.094*	-0.130**	0.036	0.094*	-0.130**	0.036	0.094*	-0.130**
	(0.020)	(0.041)	(0.042)	(0.020)	(0.041)	(0.042)	(0.020)	(0.041)	(0.042)	(0.021)	(0.041)	(0.042)
Round 16	-0.026	-0.237***	0.263***	-0.026	-0.237***	0.263***	-0.026	-0.237***	0.264***	-0.026	-0.238***	0.263***
	(0.029)	(0.045)	(0.040)	(0.029)	(0.045)	(0.040)	(0.029)	(0.045)	(0.040)	(0.028)	(0.045)	(0.040)
Round 21	0.009	-0.341***	0.332***	0.009	-0.341***	0.332***	0.009	-0.341***	0.332***	0.009	-0.341***	0.332***
	(0.025)	(0.049)	(0.043)	(0.025)	(0.049)	(0.043)	(0.025)	(0.049)	(0.043)	(0.025)	(0.049)	(0.043)
Controls		N			Married		Married, Owns business, tax ID			Married, Owns business, tax ID		
District Fixed Effects		N			N			N			Y	
Observations	3925	3925	3925	3925	3925	3925	3925	3925	3925	3925	3925	3925

[1/] This is a dummy that takes the value of 1, if the MSE chose a hypothetical 100PEN for certain, instead of a 50/50 chance to earn 200PEN or 0PEN.

[2/] The omitted category is drawing a “good client” (green ball).

Robust standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.