Universidad del Rosario
Facultad de Economía

Job migration in a rivalry setting

Robert Fletcher
Santiago Saavedra

SERIE DOCUMENTOS DE TRABAJO

No. 235

Octubre de 2019

Job migration in a rivalry setting [*]

Robert Fletcher[†] Santiago Saavedra[‡]

August 20, 2019

Abstract

The importance of social networks in job search and migration have been well documented. However, spreading information too widely throughout networks when opportunities arise can easily lead to the tragedy of the commons – too many people depleting a limited opportunity can mean no one benefits in the end. Hence, despite the generally positive value of large social networks, we should expect the strategic sharing of information within networks. To better understand this, we study the co-migration decisions of social connections through the movements of gold miners in Colombia. In this setting, we document three facts that are easily interpretable with a model of referrals and scarce resources. First, while working with close social connections is associated with higher production, having too many miners present is ultimately associated with lower production. Second, in line with the first result, we find that more productive miners, for whom depletion of resources is a greater concern, invite fewer social connections. Finally, the connections that miners are willing to invite are heavily selected; miners tend to invite productive over non-productive peers.

JEL classification: O15, L14

---

[*] We are grateful to Edinson Bolaños, Arun Chandrasekhar, Pascaline Dupas, Matthew Jackson, Melanie Morten and Kaivan Munshi and seminar participants at PacDev 2019, ITAM, Development Economics and Policy 2018, International Labor Organization, Stanford University, Universidad de los Andes, and Universidad del Rosario.

[†] Stanford University Department of Economics, *e-mail :* `rsfletch@stanford.edu`.

[‡] Universidad del Rosario Department of Economics, *e-mail:* `santisaap@gmail.com`.

1 Introduction

Many workers hear about job opportunities through social networks (Calvó-Armengol & Jackson, 2004), and in general, the literature has shown that networks affect labor market outcomes positively (Battisti, Peri, & Romiti, 2016; L. Beaman & Magruder, 2012; Ioannides & Datcher Loury, 2004). However, these studies have largely been in a non-rival setting. For example, garment workers may be happy to pass on a job opening to a friend because both will receive their wages independently. By contrast in a rival setting, like gold mining, bringing an extra friend to a mine could cause congestion and deplete the resource faster. In this paper, we study production, migration and the arrival of friends in the context of gold miners to assess whether miners face congestion and selectively invite social connections to limit congestion. Specifically, we answer three questions: (1) Do workers produce more when working with some peers but face congestion trade-offs working with many? (2) If they do, do miners restrict the number of peers they invite to join them at a mine? (3) Given that they do, is there selection in who they invite (i.e. do they select who to invite based on skill)?

In order to answer these questions, we study the production and migration of independent miners in Colombia. These miners do not have a fixed wage in a firm; their earnings are uncertain as they depend on their skill and luck finding gold. They migrate following new mineral discoveries or when the resource has been exhausted in the place they were mining. An advantage of our data is that we can separate miners' network between family miners and miners from the same hometown.

While we focus on mining, the insights and modeling from this study can be applied to other situations with scarce opportunities. For instance, startup entrepreneurs and movie producers are eager to hire lots of talented people but must ultimately limit themselves so as to not divide up potential profits across too many people. In general, the lessons learned here apply to a large number of settings given that opportunities are frequently scarce.

Theoretically, for general production functions, we show that it is not optimal to have too many friends in rivalry settings. Intuitively, the agent is trading off the positive benefits of having friends around to boost production and the cost of having to split the gold among friends, meaning the agent is able to mine the resource for less time before it runs out. We also show for standard peer effects production functions that higher skill miners invite fewer friends to join them in a new mine.

To test the theory, we use data on 133,465 precious metals transactions by 27,116 miners in 238 municipalities. For each transaction, we have information on the date, miner, location, and quantity extracted. We have two measures of the network of a miner: (i) "family miners": identified as those with a common last name and from the same hometown (ii) "co-muni miners": from the same hometown, identified as those registered to vote on the same municipality. We assume that if they are both miners and vote in the same municipality, they know each other. Given that we rely on reported transactions, the main limitation of the data is that we cannot tell apart whether a miner is mining and not finding gold or has given up and is working in a new occupation.

Empirically, we start by showing that a miner mines more gold when they have more family members joining them, but that the benefit tapers off as more family members join and congestion is higher. There is a strictly negative effect to having miners from other municipalities at the same location, again strongly indicative of congestion.

Regarding our second question, we show that there is significant evidence that miners limit invitations to their social network connections. In particular, more productive miners, for whom depletion of resources is a greater concern, are followed by fewer family and co-municipality miners when they migrate. For these miners, the positive peer effect is offset by a greater concern that the resource will be depleted faster. This is consistent with miners limiting their invitations given out in response to the scarcity of resources.

Finally, we illustrate that the miners who follow a migrant to a new mine are relatively more productive than those that stay behind. This is consistent with miners being selective over who they invite. If they were to invite indiscriminately, we should not see this positive selection and indeed likely see negative selection as more productive miners have greater opportunity costs to migrating. Taken as a whole, we indeed find strong evidence of congestion and the limited sharing of opportunities to selected individuals.

In the last part of the paper, we simulate referrals and migration of miners under different production functions. We replicate the facts that working with social network connections can increase productivity, that more productive miners are followed by fewer friends after a migration, and that there is positive selection of co-migrants.

Together, the theory, empirics, and simulations add a unique facet to the social networks in labor markets' literature. There is already an extensive literature showing how networks affect labor market outcomes (Battisti et al., 2016; L. Beaman & Magruder, 2012; Ioannides & Datcher Loury, 2004; Bewley, 1999) and the positive effect of network size on earnings (Damm, 2009; Edin, Fredriksson, & Åslund, 2003). However, we study the less well understood effect of networks in a context of limited job opportunities or resources. L. A. Beaman (2011), the study closest to our own, studied congestion in referrals among refugees by looking at the size of social networks as proxied by the number of refugees from the same origin placed in the same area. We contribute to this line of research by studying congestion on production as well as referrals. In addition, we observe not only the place of origin, but also the family to look at finer levels of the network. We also observe repeated migrations, and how these depend on the skill of network members. This is an advantage over the one-shot decision of previous studies with international migrants.

The remainder of the paper is organized as follows: Section 2 presents the theoretical framework. Section 3 describes the data, and Section 4 presents the main empirical results. Section 5 presents the results of the simulations, and the final section concludes.

2 Theoretical Framework

2.1 Setup

We present a simple model that captures the fundamental trade-off agents face: increased productivity from working with friends and having to split the output among those friends. This model will be a continuous time model from the perspective of a single agent who is managing a single mine. Managing a mine in this example consists of deciding N, the number of other miners to invite to the mine. To describe productivities, we will let the main agent have a productivity coefficient λ and assume that all friends have identical productivities λ_f. All agents produce according to the function $F(\lambda_i, \lambda_{-i})$ where $\lambda_{-i} = \sum_{j \neq i} \lambda_j$.

To start with, this agent maximizes lifetime utility

$$U(s, N) = \int_{t=0}^{\infty} e^{-\rho t} u(t)$$

where s is the stock of gold in the mine at time t_0 when the miner begins. While the mine is running and the stock of gold is still positive, $u(t)$ is simply the production the agent is getting from the mine. After the mine has run out, the agent goes to an alternative source of income that pays B. Succinctly, while the mine still has gold in it, $u = F(\lambda, N\lambda_f)$. After the mine runs out $u = B$. Assume that B is small enough such that the miner will always prefer to mine while there is gold remaining. Hence, we can write the agents lifetime utility as

$$U(s, N) = \frac{F(\lambda, N\lambda_f)}{\rho} + e^{-\rho T(N)} \left(\frac{\beta - F(\lambda, N\lambda_f)}{\rho} \right) \tag{1}$$

Note that the time it takes for a mine to run out, $T(N)$, depends on how many miners are – intuitively having a greater or smaller number of miners affects how quickly the mine is depleted. Mathematically, $T(N)$ is the stock divided by how quickly the miners collectively deplete the stock.

$$T(s,N) = \frac{s}{\sum_j F(\lambda_j, \lambda_{-j})} = \frac{s}{F(\lambda, N\lambda_f) + NF(\lambda_f, (N-1)\lambda_f + \lambda)}$$

We make the very reasonable assumption that $T_N(s,N) < 0$, reflecting the fact that if there are more miners present at a mine, that mine will run out more quickly.

Note that

$$U_T(s,N) = e^{-\rho T(s,N)}(F(\lambda, N\lambda_f) - \beta)$$

Since $\beta < F(\lambda, N\lambda_f)$, $U_T(s,N)$ is positive and agents will always benefit from having mining for longer, holding production speed constant. Also note that

$$U_F(s,N) = \frac{1}{\rho}\left(1 - e^{-\rho T(s,N)}\right)$$

which is also positive and reflects the fact that agents who increase production, while mining the same length of time, are better off. If we now look at the tradeoffs for choosing N, we have

$$U_N(s,N) = \underbrace{F_N(\lambda, N\lambda_f)}_{+} \underbrace{U_F}_{+} + \underbrace{T_N(s,N)}_{-} \underbrace{U_T}_{+} \tag{2}$$

Intuitively, the agent is trading off the positive benefits of having friends around to boost production and the cost of having to split all the gold up among friends meaning the agent is able to mine the mine for less time before it runs out.

2.2 Comparative Statics

Using the model from the previous sub-section we look at comparative statics to answer the questions we posed earlier: do workers produce more having more peers present, do more productive miners invite more or fewer peers when they move to a new location, and do they select who to inform based on skill?

In order to simplify the problem for the purposes of this version of the paper, we will assume a more specific functional for the production function. Namely, borrowing from the peer effects literature,

$$F(\lambda_i, \lambda_{-i}) = \lambda_i + \eta\lambda_{-i} + \phi\lambda_i\lambda_{-i}$$

The parameters η and ϕ capture the degree to which individuals benefit from pure returns to scale or pure peer effects respectively.Given our assumption that all the friends have the same productivity coefficient, λ_f, we can dramatically simplify the production functions. The production function for agent i simplifies to

$$F(\lambda, \lambda_{-i}) = \lambda + (\eta + \phi\lambda)N\lambda_f$$

where N is the number of friends working with agent i. The friends will produce

$$F(\lambda_f, \lambda_{-f}) = \lambda_f + (\eta + \phi\lambda)((N-1)\lambda_f + \lambda)$$

Having specified this form of the production function, we can now calculate utility functions for various sets of parameters in Matlab. In the next three subsections, we show comparative statics trends that are generally true for a variety of parameters and which in turn respond to the three questions we posed earlier.

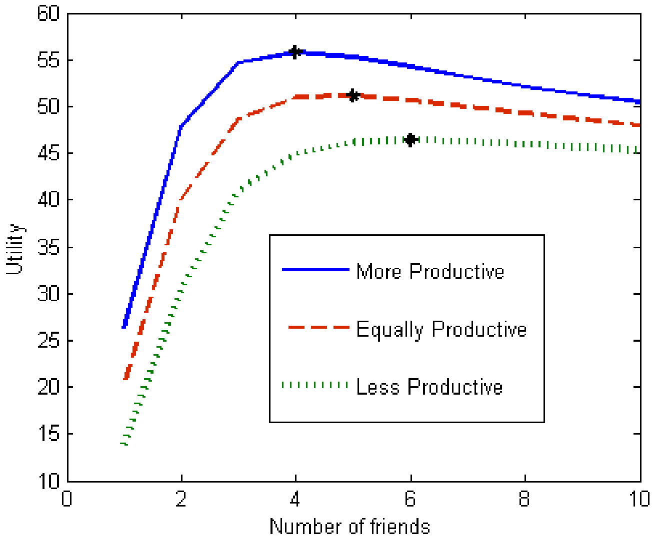
2.3 Question 1: The Cost of Too Many Friends

The first question ask whether workers produce more when they have more friends present. The answer to this question is twofold: agents do produce more at a given time while mining because friends help the miner to produce more instantaneously, however in the long run they produce less because they have to split the ore amongst their peers. This is best reflected in equation (2).

With the production function form we've just assumed, we can easily plot this tradeoff. Figure 1, which looks at various values for the miner productivity parameter, shows the concave relationship between number of friends and lifetime utility. For small numbers of friends, miners benefit substantially from inviting partners. However, the return to inviting more friends quickly slows down and eventually becomes negative. Hence, as is not captured in many models of social networks, agents have incentives to limiting their information and opportunity sharing. If miners invite all their friends, they are susceptible to the tragedy of the commons where a good thing is split too many ways.

2.4 Question 2: Agent Skill Level Comparative Statics

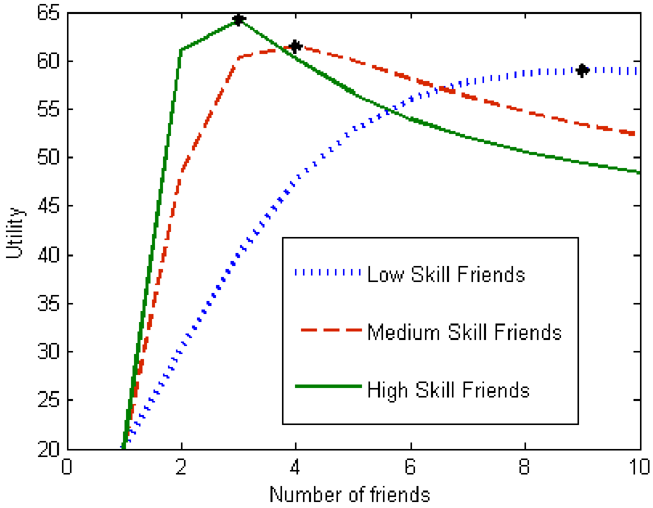
Figure 1: Optimal Selection of Partners Given Productivity of Miner



Turning now to the second question, we ask whether more productive miners invite more or fewer peers when they move to a new location. In particular, are more productive miners more self-reliant and so able to avoid having to invite too many friends. Figure 1 explores this relationship by varying the miner's productivity coefficient while holding everything else constant. Here we see that the lifetime utility curves of more productive miners are shifted upwards, which makes sense given that they are able to mine quicker and more efficiently with any level of partners. More importantly, we see that the optimal value of friends, demarcated with the mark at the optimum of each curve, decreases with the productivity of the miner. Indeed, more productive miners are less reliant on inviting partners and so don't need to split up the gold available at the mine.

2.5 Question 3: Friend Skill Level Comparative Statics

Figure 2: Optimal Selection of Partners Given Productivity of Social Connections



Last of all, we can look at which sorts of friends miner choose to invite, and whether having to limit the number of connections they invite leads them to be selective in their choices. First of all note the obvious which is that the less productive the friends are, the more of them that the agent needs to invite to reach optimal utility. Beyond this however, note that the peak of the utility function is higher the greater the productivity of the friends. Clearly, the agent is best off if they can recruit a smaller number of high skill friends. Hence, we have a very selective exclusivity that emerges – agents not only limit who they invite, but they limit it to a small number of excellent social connections.

3 Data

We have information on every reported precious metal transaction in Colombia for the years 2007-2015. For each transaction we have the date, name and national identification number of the miner ("cedula"), municipality of the transaction, type of mineral, quantity extracted and amount of royalty tax paid. The mineral can be gold, silver or platinum. The amount of tax paid is 5% of the value of the production, so we can infer the price of each transaction.[1]. With the national ID, we obtain the voting station of the miner from the National Registry, what we interpret as his municipality of origin and use as the measure of his network. In Colombia citizens get two last names: the first one the first last name of the father; the second one, the first last name of the mother. Consequently cousins share one of their two last names.

Table 1 presents summary statistics of the transactions. There are 133,465 transactions of 27,116 miners in 238 municipalities. For every transaction we can calculate the number of miners present in that municipality at the same time. We define as present another miner that registered a transaction in that municipality in a 3 months window.[2] In the last two rows of Table 1 we calculate the percentage of the miners that were working with a given miner that move with him to a new municipality. Note that it is more likely that miners from the same municipality move together in the next migration. Although one might think that miners have a cyclical pattern of migration like

---

[1] The rate is 6% for alluvial gold, 5% for platinum and 4% for silver and non-alluvial gold.

[2] The reasoning for fixing the length of the window and not using the period the miner is in the municipality is that miners that sell less often will have more miners present.

in agriculture, we show in Figure A.3 in the Appendix that this is not the case. One of the main limitations of our data is that we cannot observe whether a miner stops mining or he is mining but not reporting production.

Table 1: Summary statistics transactions

	Mean	Median	Std. Dev.	Min	Max
N family miners present	29	3	80	0	977
N co-muni miners present	253	67	428	0	3,531
N other miners present	1,155	578	1,211	0	4,587
Days since last sell	108	34	208	1	5,071
Gold grams/day	3.4	1.3	4.4	0	14
Previous rel. production co-muni	1	.98	.51	0	2.4
% family moving after	11	0	22	0	100
% co-muni moving after	11	2.8	19	0	100
% other moving after	7.8	2.6	12	0	100

Notes: An observation is a transaction. The number of co-municipality miners present is calculated as those miners that vote in the same municipality and sold gold in the same mining municipality in a three month window around the given transaction.

With the national ID we can match some of the miners to the poverty census (Sisben), where we have some socioeconomic information. Table 2 presents some characteristics of the miners. We match more than half the miners, and for most of them, the municipality of the poverty census coincides with the voting municipality. This fact confirms that is a good proxy for municipality of origin of the miners. Interestingly around half of them are women. The miners have around forty years and a low level of education.

Table 2: Summary statistics miners

	Mean	Median	Std. Dev.	Min	Max	N
In Sisben	.62	1	.49	0	1	25,930
Voting = Sisben muni	.85	1	.35	0	1	16,069
Male	.51	1	.5	0	1	16,069
Birth year	1972	1973	12	1936	1997	15,862
Education years	5.8	5	3.9	0	11	15,131
Poverty base score	30	27	19	.56	85	16,069
Monthly income	140	60	191	0	1,050	16,069

Notes: An observation is a miner. Sisben is the census of poor households for social programs in Colombia. Poverty score cutoff for cash transfers is 30. Monthly income in thousand COP 2010. Exchange rate 1USD=2,500 COP.

4 Empirical results

We answer empirically our three questions in this section. First, we show that a miner mines more gold when they have more family members joining them, but that the benefit tapers off as more family members join and congestion is higher. There is a strictly negative effect to having miners from other municipalities at the same location, again strongly indicative of congestion. Then we show that more productive miners are followed by fewer co-municipality miners in the next migration. Finally we illustrate that the miners that move are relatively more productive.

4.1 Empirical strategy

We want to estimate the effect of having more miners from the same municipality on miners production. We rely on variation from observing the same miner at different times and places. Our estimating equation is:

$$Y_{ipmt} = \beta Network_{ipmt} + \gamma_i + \gamma_m + \gamma_{Year(t)} + \varepsilon_{ipmt} \quad (3)$$

where Y_{ipmt} is output or "wage" of miner i, from voting station p, mining on municipality m on day t; $Network_{ipmt}$ indicates number of family miners or hometown miners mining at the same time in m.[3] γ_i, γ_m and $\gamma_{Year(t)}$ are miner, mining municipality and fixed effects respectively. On the appendix we will perform robustness check when using municipality-year and miner-year fixed effects to capture new discoveries on a municipality and learning/ageing of a miner. In the previous literature on non-rivalness setting $\beta > 0$. However, given the finiteness of mineral resources we expect a negative relationship.

4.2 Results

Question 1

The results of estimating equation 3 are presented in Table 3. Column 1 shows that a miner produces more when they have more family miners present. On the other hand, there is a negative correlation with having larger numbers of co-muni miners present. As miners collaborate with only a fraction of non-family miners from the same municipality, most will add to congestion but not contribute to production. Hence it is expected that the congestion effects to be larger relative to the production benefits and the overall effect to therefore be negative.

We explore this negative coefficient further in Column 2 by looking at the number of miners from other municipalities present. As expected, the negative relationship is greater for miners from other municipalities than from the same municipality as they are even more likely to only be competitors who add to congestion but do not help with production.

In Column 3 we assess the concavity of the results in Column 2. In particular, we include a separate coefficients for cases with more/less than 25 miners of each type present. We again see that working with more family members is associated with higher levels of production, however this effect tapers off as the number of family members grows. Each coefficient is significant at 10%

---

[3] In our main specification we define "working at the same time" as reporting a sell in the three month window around time t.

level. This is consistent with increased levels of congestion among even family members. We do not observe negative returns to any levels of family members, which is consistent with miners being able to limit the number of family members invited and so avoid inviting family members when the congestion effect outweighs any beneficial effects on production.

As far as miners from the same municipality, we see a slightly different relationship. We again see having miners from the same municipality is negatively associated with production, but the coefficient grows larger with the number of miners from the same municipality present. Statistically, we cannot reject that when a miner has a few co-municipality miners present the effect on production is null. But when there are too many co-municipality or other type of miners the effect is negative and significant. Seeing as the fraction of these miners who are collaborating as opposed to competing likely grows with the number present, this again is consistent with congestion from having more miners present. Last of all, we see a strictly negative relationship from having more miners from other municipalities present. Note that regardless of the quantity, the association is positive for family, negative for co-municipality miners, and the most negative for other miners. This strongly suggests that our network definition is indeed capturing closeness.

We perform a battery of robustness of the main results on Tables A.1-A.5. Table A.1 repeats the specifications of Table 3, but adding municipality-year fixed effects on Columns (1) - (3) and also miner-year fixed effect, Columns (4)-(6). Understandanbly some coefficients are not longer significant, but the key message that a few family friends increase production but too many other miners crowd out is still present. Table A.4 presents log-log and log-linear specifications: results using logarithm of production as the dependent variable and logarithm of friends as the independent. The significance of the coefficients is maintained, although for overall family miners present it is not significant. On Columns (3) and (6) when we separate by small and large number of friends present, they are both significant. Table A.3 add more than two bins of family friends, and Table A.2 uses other cut for small and large number of family miners. Finally, we exclude every transactions from gold buyers found to be corrupt. Table A.5, shows that we lose a sixth of observations but the coefficients are basically unaffected. In fact in Column 3, now the coefficient of small number of other miners present is now significant.

Table 3: Effect of more friends on own production

Dependent variable:	Gold sold (grams)		
	(1)	(2)	(3)
N family miners present	0.083**	0.082**	
	(0.039)	(0.039)	
N co-muni miners present	-0.016***	-0.012***	
	(0.0051)	(0.0047)	
N other miners present		-0.0091***	
		(0.0031)	
N family present * (< 25)			0.33***
			(0.082)
N family present * ($>= 25$)			0.089**
			(0.039)
N co-municipality miners present * (< 25)			-0.079
			(0.082)
N co-municipality miners present * ($>= 25$)			-0.014***
			(0.0047)
N other miners present * (< 25)			-0.34
			(0.22)
N other miners present * ($>= 25$)			-0.0093***
			(0.0031)
N. of obs.	133,465	133,465	133,465
Municipalities	238	238	238
Mean of Dep. Var.	75.8	75.8	75.8
R^2	0.57	0.57	0.57

Notes: Regressions estimating equation 3 using the transaction data. All regressions include miner, mining municipality and year fixed effects. Standard errors, clustered by municipality, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Question 2

In Table 4 we estimate whether the fraction of family and co-municipality miners that follow a miner depend on the relative skill of the miner. In order to test this, we calculate for each migration the fraction of family and co-municipality miners working at the same time in the origin municipality that moved to the destination municipality after miner i. Our measure of skill of miner i is his production in the origin municipality divided by the average production of his family or co-municipality miners in the origin municipality. Column 1 shows that a miner that was more productive is followed by fewer co-workers. This is in line with the theoretical framework. In Column 2 we test whether the relationship depends on the number of friends, and find that irrespectively on the number of friends high skill miners have fewer co-municipality miners following them after a migration.

Question 3

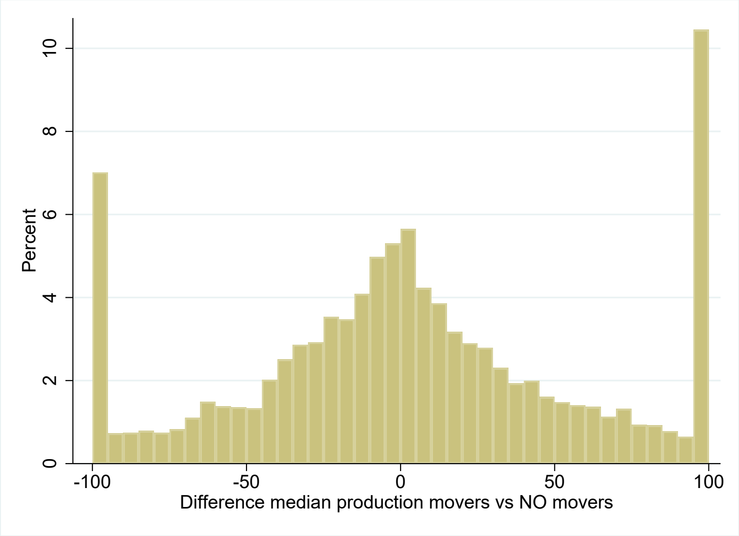
Finally we test whether the co-municipality workers that follow a miner after a migration are more productive. For each miner migration we calculate the median production before the move of co-municipality miners that followed the mine. We also calculate the median production for those co-municipality miners that did not move. Figure 3 plots the histogram of the differences of these two medians. We observe that the distribution is concentrated to the right, suggesting that movers are more productive.

5 Simulation results

In this section we present the equivalent regressions to our main empirical specifications, but using data generated with model simulations. We show that the simulations do not fit the empirical facts when miners are unable to limit the spread of information on mining opportunities. Once miners can limit the number of inviations extended to other miners, we find that the model fits the empirical results.

Consider the decision of a miner that can stay mining where he is or can migrate to a place with better prospect. He makes the decision based on the information he has through friends about the mining stock in other munici-

Figure 3: Histogram with differences between movers and not movers production



Notes: Histogram of the differences between the median production of miners that move after a miner migration and those that stay at the origin municipality. Extreme values bunched for visualization.

palities. The matrix $\mathbb{I}_{N\times M}$ has entries $I[i,m]=1$, if miner i knows about the stock in municipality m, either because he is there or a friend informed him. If miner i invites miner j to municipality m_i, then $\mathbb{I}[j,m_i]=1$. After receiving the invitations a miner stays or migrate to m^*, where

$$m^* = \arg\max\left\{\max_{m:\mathbb{I}[i,m]=1} V(i,m), \mathbb{E}_{m:\mathbb{I}[i,m]=0}V(i,m)\right\}$$

Where $V(i,m)$ is the discounted value of miner i producing at municipality m. Miner i has individual productivity of λ_i. We consider two possible production functions:

- Peer congestion: $F_i(\lambda_i,\vec{\lambda}) = \lambda_i + \sum_{j,m_i=m_j}\lambda_i\lambda_j/N_i$
- Peer effects: $F_i(\lambda_i,\vec{\lambda}) = \lambda_i + \sum_{j,m_i=m_j}\lambda_i\lambda_j$

We perform simulations, using the following parameters based on the empirical data. $M=238$, mining municipalities. 544, voting municipalities i.e. home municipalities of the miners, where they know their friends. $N = 25,930$ number of miners. And we run the simulation for $T = 10$ periods. We draw each miners productivity from the uniform distribution ($\lambda \sim U(0,1)$). Each miner is from a random voting municipality, consequently the friends are random. And each miner starts mining in random municipality, and then start moving optimally based on the information he has. We consider two possible ways of doing migration invitations. (A) Core equilibrium: miners can invite as many friends as they want, but if they "experience" there is much congestion on the new location they try a "new" one until nobody wants to try to move. Computationally miners move, but we do not count a production period until every miner does not have an incentive to move. This algorithm is based on based on (Kelso Jr & Crawford, 1982). (B) Pair match: each miner is randomly matched with one of his friends, and they decide whether to invite each other. That is, a miner can invite maximum one miner each period.

Question 1

The results of estimating equation 3 with the simulated data are presented in Table 5. Column 1 presents the results with the "Peer congestion" production function and Column 2 with the traditional "peer effects". Panel A presents the result with the "Core equilibrium" invitations and Panel B: with the "Pair match" algorithm. On the "Core equilibrium" a miner can invite many miners on his network. Fittingly, we see a negative relationship in Column 1 from having more co-municipality members present. The classic tragedy of the commons problem emerges, and there is overcrowding. Having more network miners present is associated with negative effects on production. It is not until we allow miners to limit the number of invitations extended that we observe strictly positive returns to having other family members present, as observed in Table 3 with the actual data.

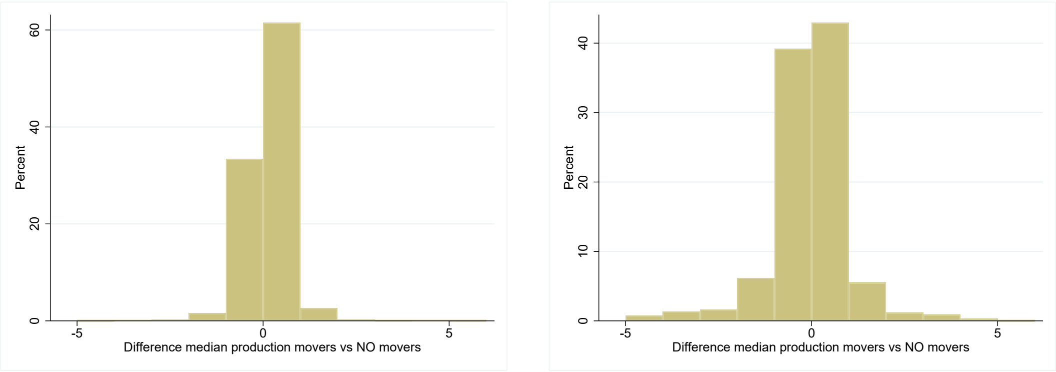
Question 2

Table 6 is the analogous of Table 4. Similar to Table 5 each column contains the results for a different production function, and each panel for a different way of simulating invitations. We replicate the empirical results that more productive miners are followed by fewer co-municipality workers for the "Core equilibrium" simulation regardless of the production function.

Question 3

Finally we test with the simulated data whether the co-municipality workers that follow a miner after a migration are more productive. Figure 4 is the analogous of 3 with the simulated data. We observe that for both functions the distribution is concentrated to the right, suggesting that movers are more productive. Note that the selection is more clear for the peer effects function.

Figure 4: Histogram with simulated differences between movers and not movers production



Notes: Histograms of the simulated differences between the median production of miners that move after a miner migration and those that stay at the origin municipality. The left figure is using the congestion production function and the right histogram using the traditional peer effects function.

6 Conclusions

In this paper, we study job migration and networks in the case of gold miners in Colombia. As the mining resource is finite, we document evidence of congestion. Furthermore, we show that miners appear to respond to congestion effects by limiting the number of other miners they invite to work with them, even though working with other miners can have positive effects on production. We also show that evidence of selection when inviting other miners to join, in the form of selection on the productivity of miners. We interpret these results as strong evidence of congestion and potentially negative effects from sharing opportunities too widely across social networks.

References

Battisti, M., Peri, G., & Romiti, A. (2016, July). *Dynamic effects of co-ethnic networks on immigrants' economic success* (Working Paper No. 22389). National Bureau of Economic Research. Retrieved from `http://www.nber.org/papers/w22389` doi: 10.3386/w22389

Beaman, L., & Magruder, J. (2012, December). Who gets the job referral? evidence from a social networks experiment. *American Economic Review*, *102*(7), 3574-93. Retrieved from `http://www.aeaweb.org/articles?id=10.1257/aer.102.7.3574` doi: 10.1257/aer.102.7.3574

Beaman, L. A. (2011). Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the us. *The Review of Economic Studies*, *79*(1), 128–161.

Bewley, T. (1999). *Why don't wages fall in a recession.* Harvard University Press Cambridge.

Calvó-Armengol, A., & Jackson, M. O. (2004, June). The Effects of Social Networks on Employment and Inequality. *American Economic Review*, *94*(3), 426-454.

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, *27*(2), 281-314. Retrieved from `http://www.jstor.org/stable/10.1086/599336`

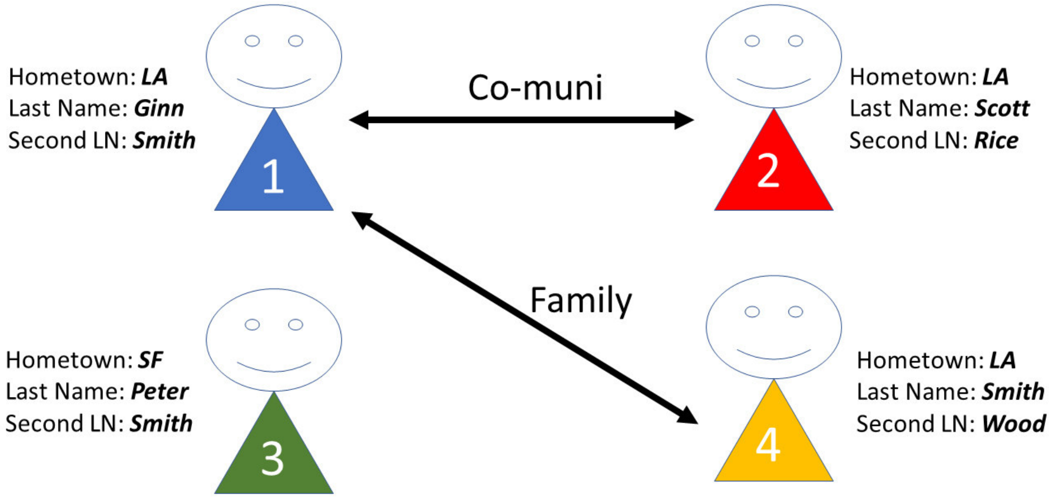
Edin, P.-A., Fredriksson, P., & Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment*. *The Quarterly Journal of Economics*, *118*(1), 329-357. Retrieved from `+http://dx.doi.org/10.1162/00335530360535225` doi: 10.1162/00335530360535225

Ioannides, Y. M., & Datcher Loury, L. (2004, December). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, *42*(4), 1056-1093. Retrieved from `http://www.aeaweb.org/articles?id=10.1257/0022051043004595` doi: 10.1257/0022051043004595

Kelso Jr, A. S., & Crawford, V. P. (1982). Job matching, coalition formation, and gross substitutes. *Econometrica: Journal of the Econometric Society*, 1483–1504.

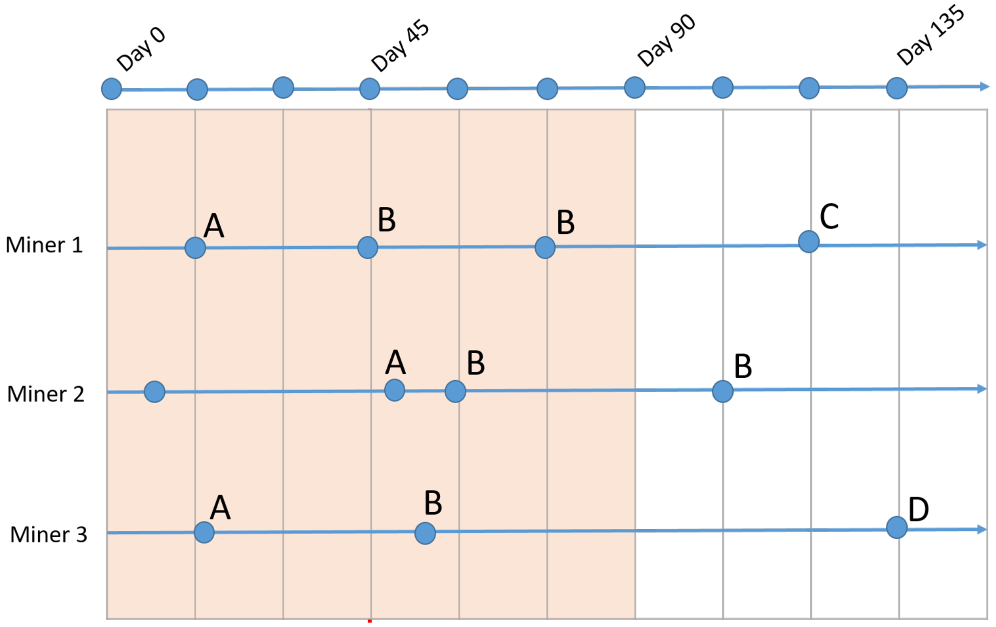
Appendix A Additional Figures and Tables

Figure A.1: Network definition



Notes: Example of four miners and their connections. Miner 1 and miner 2 are co-muni miners because they are from the same hometown, but do not share a last name. Miner 3 is from SF so is not connected to Miner 1. Miner 4 is family of miner 1, because they are from the same hometown and have a last name in common "Smith". Miners 2 and 4 are co-muni miners, but the arrow is not shown.

Figure A.2: Definition of a friend moving after



Notes: Example of definition of a friend moving after. Each row indicates the registered transactions of a miner. All miners have transactions on municipality A and then on B. We define that miner 2 moved after miner 1 from A to B, because after miner 1 registered a transaction on B, miner 2 had a transaction on A. However, we do not say miner 3 moved after miner 1, because 3 did not sell in another municipality while 1 was on A. They could have move together and sell in different days.

Table 4: Fraction of family/co-muni that follow a miner depending on skill

Dependent variable:	Fraction of network miners moved	
	(1)	(2)
Panel A: Family		
Previous rel. production family	-0.13***	
	(0.040)	
Previous rel. production * family $< p(50)$		-0.097**
		(0.044)
Previous rel. production * family $>= p(50)$		-0.097
		(0.089)
N. of obs.	35,911	35,774
Municipalities	85	83
Mean of Dep. Var.	11.7	11.3
R^2	0.46	0.58
Panel B: Co-muni		
Previous rel. production co-muni	-1.06**	
	(0.41)	
Previous rel. production * friends $< p(50)$		-1.61***
		(0.45)
Previous rel. production * friends $>= p(50)$		-0.41
		(0.56)
N. of obs.	42,860	42,860
Municipalities	95	95
Mean of Dep. Var.	11.6	11.6
R^2	0.55	0.55

Notes: All regressions include miner, mining municipality and year fixed effects. Standard errors, clustered by municipality, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Simulation effect of more friends on own production

Dependent variable:	Gold sold (grams)	
	Congestion (1)	Peer effects (2)
Panel A: Core equilibrium		
N co-municipality miners present	-0.0089***	0.0099*
	(0.0019)	(0.0052)
N. of obs.	155,580	155,580
Municipalities	238	238
Mean of Dep. Var.	0.74	0.81
R^2	0.89	0.68
Panel B: Pair match		
N co-municipality miners present	0.012***	0.11***
	(0.0019)	(0.010)
N. of obs.	155,580	155,580
Municipalities	238	238
Mean of Dep. Var.	0.77	1.04
R^2	0.94	0.72

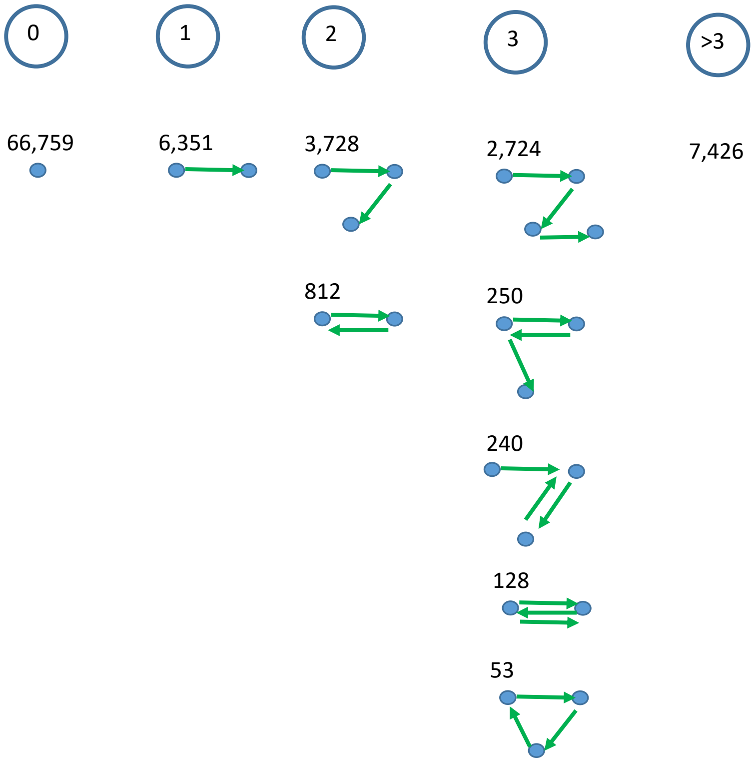
Notes: Regressions estimating equation 3 using simulated data. All regressions include miner, mining municipality and year fixed effects. Standard errors, clustered by municipality, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Simulation fraction of friends that follow a miner depending on skill

Dependent variable:	Fraction of co-municipality miners moved	
	Congestion	Peer effects
	(1)	(2)
Panel A: Core equilibrium		
Previous rel. production	-0.0018	-0.000028
	(0.0015)	(0.000090)
N. of obs.	21,095	29,618
Municipalities	201	85
Mean of Dep. Var.	3.48	5.69
R^2	0.44	0.62
Panel B: Pair match		
Previous rel. production	0.00097	0.00020
	(0.0070)	(0.00013)
N. of obs.	836	39,444
Municipalities	200	238
Mean of Dep. Var.	0.29	0.60
R^2	0.73	0.49

Notes: Standard errors, clustered by municipality, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A.3: Migration patterns



Notes: Number of miners in each migration pattern by number of migrations.

Table A.1: Robustness: Effect of more friends on production with mining municipality-year and miner-year fixed effects

Dependent variable:	Gold sold (grams)					
	(1)	(2)	(3)	(4)	(5)	(6)
N family miners present	0.053	0.054		0.058	0.058	
	(0.039)	(0.039)		(0.050)	(0.050)	
N co-muni miners present	-0.011**	-0.011**		-0.015*	-0.014*	
	(0.0052)	(0.0053)		(0.0079)	(0.0074)	
N other miners present		-0.0056*			-0.0059	
		(0.0033)			(0.0036)	
N family present * (< 25)			0.20***			0.17*
			(0.078)			(0.096)
N family present * ($>= 25$)			0.058			0.059
			(0.039)			(0.051)
N co-municipality miners present * (< 25)			-0.11			-0.18
			(0.071)			(0.12)
N co-municipality miners present * ($>= 25$)			-0.012**			-0.014*
			(0.0052)			(0.0074)
N other miners present * (< 25)			0.042			0.11
			(0.21)			(0.25)
N other miners present * ($>= 25$)			-0.0057*			-0.0060*
			(0.0033)			(0.0035)
Fixed Effects	MY	MY	MY	MYIY	MYIY	MYIY
N. of obs.	133,250	133,250	133,250	118,432	118,432	118,432
Municipalities	231	231	231	229	229	229
Mean of Dep. Var.	75.8	75.8	75.8	72.8	72.8	72.8
R^2	0.60	0.61	0.61	0.68	0.68	0.68

Notes: This Table replicates the results of Table 3 including mining municipality-year (MY) and miner-year (IY) fixed effects. Columns (1)-(3) include MY and Columns (4)-(6) MY and IY. Regressions estimating equation 3 using the transaction data. Standard errors, clustered by municipality, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Robustness: Effect of more friends on log production and production per day

Dependent variable:	Log(Gold sold (grams))			Gold mined per day (grams/day)		
	(1)	(2)	(3)	(4)	(5)	(6)
N family miners present	0.0013	0.0013		0.0020	0.0019	
	(0.00096)	(0.00097)		(0.0015)	(0.0014)	
N co-muni miners present	-0.00036***	-0.00031***		-0.00075*	-0.00065*	
	(0.000086)	(0.000085)		(0.00043)	(0.00037)	
N other miners present		-0.00013***			-0.00021	
		(0.000047)			(0.00018)	
N family present * (< 25)			0.0071***			0.019***
			(0.0020)			(0.0058)
N family present * ($>= 25$)			0.0015			0.0025*
			(0.00097)			(0.0014)
N co-municipality miners present * (< 25)			-0.00085			-0.0013
			(0.0016)			(0.0057)
N co-municipality miners present * ($>= 25$)			-0.00034***			-0.00073**
			(0.000085)			(0.00037)
N other miners present * (< 25)			-0.0090**			-0.0042
			(0.0044)			(0.012)
N other miners present * ($>= 25$)			-0.00013***			-0.00022
			(0.000047)			(0.00017)
N. of obs.	133,465	133,465	133,465	100,306	100,306	100,306
Municipalities	238	238	238	215	215	215
Mean of Dep. Var.	3.77	3.77	3.77	3.41	3.41	3.41
R^2	0.62	0.62	0.62	0.41	0.41	0.41

Notes: This Table replicates the results of Table 3 but using as depent variable logarithm of output (Clumns (1)-(3)) or output per day (Columns (4)-(6)). Regressions estimating equation 3 using the transaction data. All regressions include miner, year and mining municipality fixed effects. Standard errors, clustered by municipality, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Robustness: Non-parametric family network size

Dependent variable:	Gold sold (grams)		
	(1)	(2)	(3)
Bin size k:	5	10	25
N family > 0	3.81***	3.96***	3.75***
	(0.94)	(0.96)	(0.95)
N family $> k$	3.61***	4.22***	5.87***
	(1.26)	(1.59)	(2.04)
N family $> 2k$	1.95*	3.29*	-2.38
	(1.07)	(1.86)	(3.42)
N family $> 3k$	2.07	6.69	3.11
	(1.45)	(4.19)	(3.46)
N family $> 4k$	2.75	-8.16	0.31
	(2.14)	(5.42)	(3.82)
N. of obs.	133,465	133,465	133,465
Municipalities	238	238	238
Mean of Dep. Var.	75.8	75.8	75.8
R^2	0.57	0.57	0.57

Notes: Regressions estimating equation 3 non-parametrically using the transaction data. All regressions include miner, year and mining municipality fixed effects. Standard errors, clustered by municipality, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.1 Definition of a valid transaction

There are concerns of misreporting and money laundering around gold mining in Colombia. Specially because around 84% of the gold mining area is illegally mined (?, ?).For the definition of a valid transaction in our main results we used the following conditions

1. A miner cannot sell more than 1500 grams of gold in a year
2. A miner cannot be registered as dead at the time of the transaction
3. A miner cannot sell gold in more than one municipality in the same day

Table A.4: Robustness: Threshold for small/large network

Dependent variable:	Gold sold (grams)		
Threshold:		Cut	
	10 (1)	50 (2)	100 (3)
N family present * ($< cut$)	0.21	0.35***	0.18***
	(0.18)	(0.080)	(0.050)
N family present * ($>= cut$)	0.083**	0.093**	0.087**
	(0.039)	(0.038)	(0.037)
N co-muni miners present * ($< cut$)	-0.19	0.0011	-0.028
	(0.20)	(0.054)	(0.034)
N co-muni miners present * ($>= cut$)	-0.012***	-0.016***	-0.015***
	(0.0047)	(0.0052)	(0.0057)
N other miners present * ($< cut$)	-0.33	-0.28**	-0.11
	(0.47)	(0.14)	(0.084)
N other miners present * ($>= cut$)	-0.0091***	-0.0095***	-0.0092***
	(0.0031)	(0.0031)	(0.0032)
N. of obs.	133,465	133,465	133,465
Municipalities	238	238	238
Mean of Dep. Var.	75.8	75.8	75.8
R^2	0.57	0.57	0.57

Notes: This Table replicates the results of Table 3 column (3) but using alternative thresholds for small/large network. Regressions estimating equation 3 using the transaction data. All regressions include miner, year and mining municipality fixed effects. Standard errors, clustered by municipality, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A more stringent criteria is to delete all the transactions from the gold buying agencies found to be corrupt. We recalculated the number of family, co-muni miners and other miners present, and ran again the main regression. Results are presented on Table A.5

Table A.5: Effect of more friends on own production

Dependent variable:	Gold sold (grams)		
	(1)	(2)	(3)
N family miners present	0.085**	0.086**	
	(0.040)	(0.040)	
N co-muni miners present	-0.016**	-0.013**	
	(0.0070)	(0.0062)	
N other miners present		-0.0086***	
		(0.0031)	
N family present * (< 25)			0.29***
			(0.094)
N family present * ($>= 25$)			0.090**
			(0.040)
N co-muni miners present * (< 25)			-0.064
			(0.072)
N co-muni miners present * ($>= 25$)			-0.014**
			(0.0063)
N other miners present * (< 25)			-0.53**
			(0.22)
N other miners present * ($>= 25$)			-0.0088***
			(0.0031)
N. of obs.	114,570	114,570	114,570
Municipalities	213	213	213
Mean of Dep. Var.	72.9	72.9	72.9
R^2	0.59	0.59	0.59

Notes: Regressions estimating equation 3 using the alternative definition of valid transaction data. All regressions include miner, mining municipality and year fixed effects. Standard errors, clustered by municipality, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$