Endogenous Matching and the Empirical Determinants of Contract Form

Daniel A. Ackerberg and Maristella Botticini*

Current Version: May 10, 2000

Abstract

Theoretical work on contracts often identifies an "optimal" contract as a function of the characteristics of the principal and agent who are contracting. Correspondingly, empirical work on contracts often regresses contract choice on observed (by the econometrician) characteristics of the principal and agent. This paper examines the econometric implications when some of the theoretically relevant characteristics are partially observed (e.g. proxied) or unobserved. We show that if there are incentives whereby particular types of agents end up contracting (i.e. matching) with particular types of principals (and we argue that there are), one may end up with estimated coefficients on the observed characteristics that are misleading. The problem is that this *endogenous matching* generates correlation between observable characteristics of one party and proxy errors of the other party, causing biases in many or all coefficients of interest. We suggest a number of solutions to this problem and apply these solutions to a historical dataset on agricultural contracts between landlords and tenants in Renaissance Tuscany. Our major findings are: (i) controlling for endogenous matching has a significant impact on parameters of interest, and (ii) tenants' risk aversion appears to have played a role in contract choice. This is an interesting result since previous studies in both agrarian and franchising settings have not found support for the risk sharing hypothesis.

*Economics Dept., Boston University, Boston, MA 02215 and NBER (ackerber@bu.edu), Economics Dept., Boston University, Boston, MA 02215 (maristel@bu.edu). We thank John Conley, Russell Cooper, Ester Duflo, Maitreesh Ghatek, Hseuh-Ling Huynh, Sam Kortum, Kevin Lang, Bob Miller, Dilip Mookerjee, Mike Riordan, and Mike Whinston as well as seminar participants at BU, Brown, Duke, ECU, McGill, Minnesota, MIT I.O Seminar, MIT Development Seminar, Princeton, UCLA, Yale, the 1999 North American Econometric Society Meetings and the 1999 NBER Winter IO meetings for helpful comments and suggestions.

1. Introduction

Theoretical work on contract choice often starts with a principal with particular characteristics, an agent with particular characteristics, and characteristics of the task to be contracted on. One usually proceeds to solve for the optimal contract form (e.g. the share of output to be given to the agent, transfer payments to be made) as a function of these characteristics. Among other things, recent work has examined the implications of factors such as risk-aversion, monitoring ability, moral hazard, capital constraints, transactions costs, and multiple tasks on optimal contracts and the second best outcome (e.g. Cheung 1969; Stiglitz 1974; Mirrlees 1974; Holmstrom 1979; Holmstrom and Milgrom 1987, 1991, 1994; Milgrom and Roberts 1992).

Empirical work on contract choice, often in an agrarian or franchising context, usually starts with such equations in mind and proceeds by regressing contract choice on observed principal, agent, and task characteristics. The point of such works is that knowing if, how, or how much certain characteristics affect contract choice can tell us something about which of the above factors are important. This in turn can provide us with valuable information about the functioning of a micro-economy. For example, if risk sharing appears to be an important determinant of contract choice, one might make inferences about the state of insurance markets in an economy. Similarly, measuring the impact of potential capital constraints can shed light on the functioning of capital markets. Such knowledge can be beneficial for policy, particularly in a development context.

Much of the empirical contract choice literature has considered agrarian settings, where different types of contracts are often used at the same time and place. For example, crop-share, fixed rent, and wage contracts were all used in medieval Tuscany, in early modern France and Spain, and in the post-bellum U.S. South.[1] Share contracts are still common agrarian arrangements in North American agriculture, as well as in many developing countries (Allen and Lueck 1992, 1993, 1995; Cheung; Laffont and Matoussi 1995; Le- er and Rucker 1991; Rao 1971; Shaban, 1987). Business franchising, where one typically observes wide ranges of royalty rates and franchise fees, has also been studied (e.g. Brickley and Dark 1987; Martin 1988; Lafontaine 1992; Slade 1996).

Much of this empirical literature, both in agriculture and franchising, has focused on testing two possible determinants of contract choice. On one hand, risk-sharing models stress that, in the presence of a risk-averse agent who can shirk in performing the tasks assigned by the principal, share or royalty contracts offer insurance and, at the same time, provide incentives for the agent to be diligent (Cheung; Holmstrom; Holmstrom and Milgrom; Milgrom and Roberts; Otsuka and Hayami 1993; Stiglitz). On the other hand, transaction-cost models tend to ignore risk preferences and focus on enforcement costs

[1] See Alston (1981), Alston and Higgs (1982), Bloch (1970), Carmona and Simpson (1998), Emigh (1997), Galassi, Mealli, and Pudney (forthcoming), Galassi and Kauffman (1997), Herlihy and Klapisch (1978), Higgs (1974), Hoffman (1984), Kauffman (1993), Reid (1973; 1976; 1977).

and transaction-specific assets (Allen and Lueck; Eswaran and Kotwal 1985; Hallagan 1978; Lafontaine; Le- er and Rucker; Mulherin 1986). Interestingly, there seems to be little empirical support of risk sharing as an important determinant of contract choice in either franchising or agriculture. Allen and Lueck (1995, 447) state,

> Accumulated evidence confronting risk-sharing and transaction costs—covering such topics as franchising, gold mining, sharecropping, and timber—actually favors the transaction-cost framework.

Other work has found empirical support for moral hazard (Lafontaine), capital constraints (Brandt and Hosios; Laffont and Matoussi), and multitasking issues (Slade; Galassi, Mealli, and Pudney) as important determinants of contractual arrangements.

While the existing empirical literature on the determinants of contract choice is vast and interesting, we feel that there is a potential problem with much of the above work that deserves attention. A key dichotomy between the theoretical and empirical literatures is that in the theory literature, there is no measurement problem regarding principal, agent, and task characteristics. In contrast, in the empirical literature there clearly is. Many potentially relevant characteristics may be unobserved, partially observed, or observed with error by an econometrician. This observability problem is often acknowledged or mentioned in passing by the empirical literature, but the implications do not seem to have been fully discussed. This paper argues that if principals and the agents they contract with are "matched" with each other according to economic variables (and we argue that there are clearly incentives for such *endogenous matching*), this observability problem is important and casts doubt on estimated coefficients in regressions of contract choice on observed characteristics.[2] More importantly, we suggest techniques for ameliorating these problems and show that these techniques make a difference in a dataset on agrarian contracts.

To exemplify incentives for endogenous matching and its implications, consider Allen and Lueck (1992), who examine whether the inherent riskiness of a crop affects the type of contract used for that crop. Their hypothesis is that if risk-effects are an important determinant of contract choice, then higher risk crops will be more likely to utilize sharecropping rather than fixed rent contracts.[3] Empirically, they do not find this correlation and conclude that risk-sharing is not an important determinant.

Now consider an alternative explanation of this empirical result where risk-sharing *is in fact* important (e.g. the Holmstrom-Milgrom (1987) model). Suppose that half the potential tenants in the economy are risk neutral, while the other half are risk averse.

[2] While there is a large theory literature on matching (e.g. Shapley and Shubik 1972; Crawford and Knoer 1981; Becker 1981; Shimer and Smith 1996; Legros and Newman 1998), we know of no papers that address these empirical implications of matching.

[3] Their maintained assumption is that the type of crop grown on a particular plot of land is fixed (e.g. due to climate or soil type) and not a choice variable of the landlord.

Similarly, half the crops are very risky, and half are somewhat less risky. From a social welfare point of view, one can show that it would be best for the risk neutral tenants to work on the very risky crops (see section 3). If this "endogenous matching" equilibrium were exactly the outcome, the risky crops will actually be associated with fixed-rent contracts (because the tenants are risk neutral, the optimal contract is fixed rent) while the less risky crops, which are worked by risk averse tenant, would be associated with share contracts. Note that this extreme example gives *exactly the reverse* empirical implication than that argued by Allen and Lueck—fixed-rent contracts are found on the *risky* crops. The problem here is that while "riskiness of crop" may be exogenous to the landlord who owns the land, it is endogenous, through principal-agent matching, to the type of tenant attracted to it.[4]

If tenants' risk aversion was perfectly observed by the econometrician, one could solve this endogeneity problem by regressing contract choice on crop riskiness *and* risk aversion.[5] However, economists rarely profess to exactly observe a tenant's risk aversion. The story is typically that one has a proxy or proxies for risk aversion such as wealth or property, that might be used instead. Importantly, we show that using proxies for risk aversion in such regressions *does not* solve the endogenous matching problem. With endogenous matching, the crop variability variable will still be correlated with error term through the proxy error, i.e. the unobserved component of risk aversion. This correlation will directly bias the estimated coefficient on crop riskiness. It will also bias the coefficient on the risk-aversion proxy due to correlation between the proxy variable and crop riskiness.

The above example considered matching based on risk and risk aversion. There are many other stories that suggest matching between heterogeneous principals and agents. For example, principals with more ability to monitor or more ability to measure output (who might relatively prefer lower share or higher royalty contracts) might end up matching with agents with more risk aversion, more credit constraints, or a higher cost of effort (who might also relatively prefer low share or wage contracts). We argue that *any* such matching can be a serious problem when one is relying on proxies for relevant variables, which is often the case in much of the contract choice literature. Put succinctly, the matching generates correlation between observable characteristics of one of the parties and proxy errors of the other party, potentially biasing many or all coefficients of interest.

This paper addresses this problem and suggests potential solutions. These solutions follow two general approaches: both require consideration of a "matching equation" that describes how principals and agents are matched with one another. The first solution relies on instrumental variables. What we want are instruments that affect the matching

[4] One can make similar arguments in the franchising context where contract choice (e.g. royalty rate) is regressed against measures of riskiness. Again, if particular types of entrepeneurs (e.g. risk neutral or risk loving ones) are attracted to risky franchises, the results of such a regression are questionable.

[5] Although we note that one would run into collinearity problems if matching caused crop variability and risk aversion to be too closely correlated.

equation but do not affect the contractual choice or proxy equations. We suggest the use of geographical based instruments that can affect the matching equation through, e.g., differences in the exogenous distribution of land type across regions.

The second solution revolves around covariance restrictions. The idea here is that one can use correlation between observed characteristics to measure and control for correlation between observed characteristics and proxy errors. We also argue that this matching equation can not only help solve endogeneity problems in contract choice equations, but that it can also provide economically interesting information in its own right.

Lastly, we apply our techniques to historical data on contracts between landlords and tenants in Renaissance Tuscany. This is a similar dataset to that used by Galassi, Mealli, and Pudney (1999), with the significant addition of variables measuring tenant characteristics. These tenant variables are important because they allow us to both test for, and econometrically control for, endogenous matching.[6] Not only do we discover such matching, but we find that controlling for the matching can have a fairly large impact on both the magnitude and significance of contract choice coefficients. For example, *without* controlling for matching, there is little if no effect of tenant wealth on contract choice. *After* controlling for the biases due to endogenous matching, we find a stronger and significant effect of tenant wealth. Interestingly, both of our solutions support similar conclusions.

Given that wealth proxies for risk aversion, this result suggests that risk-sharing is an important determinant of contract choice.[7] This result is particularly interesting given the lack of evidence of risk-sharing in the prior empirical literature. It is also suggestive that prior work may suffer from biases because matching is ignored. In addition to correcting for endogenous matching in contract choice equations, we also make interesting economic inferences from our estimated matching equations, i.e. the way in which principals and agents match. In our dataset, the direction of matching is consistent with multitasking issues being an important factor in contract choice. Lastly, we examine effects of landlord characteristics, in particular a landlord's monitoring ability, on contract choice. While the signs of our coefficients are sensible, our results are not statistically significant.

From a historical perspective, our results provide an answer to a long-standing debate on why share contracts became a predominant agrarian arrangement in Tuscany after

[6] We examined a subset of this data in a prior paper, following the rest of the literature in ignoring potential matching between principals and agents (Ackerberg and Botticini, forthcoming).

[7] Another conceivable interpretation of the correlation between low wealth and share contracts is that of limited liability. Articles by Basu (1992) and Sengupta (1997) obtain such a result (under some conditions). However, these articles rely on tenants having a second choice variable (in addition to effort), that of choosing the riskiness of the production procedure. In essence, share contracts are used to prevent low wealth tenants from using procedures that are too risky. Our feeling is that in our particular dataset, there was not significant latitude for tenants to choose a level of riskiness (as there might be currently, e.g. by chooing a type of seed, for example: high yield and weather sensitive vs. low yield and weather insensitive).

the Black Death. Although Renaissance Tuscany was a vibrant economy, endowed with sophisticated economic institutions, her agricultural insurance, rural credit, and capital markets were either missing or imperfect, as is also the case in many contemporary developing countries. Therefore, while the historical debate is interesting in itself, the fact that we have wealth of information on landlord and tenant characteristics not easily available for contemporary countries also justifies the choice of this historical dataset. We think that our findings can also provide insights on issues currently debated in the context of many developing countries, foremost the role of contractual arrangements in substituting for missing or imperfect insurance and/or capital markets.

2. Setup

Consider a standard moral hazard model in which a principal and an agent are contracting over a task to be done. There is no hidden information, i.e. the principal and agent each know the characteristics of their contracting partner. Suppose theory gives us a contract choice equation which describes the second best contract y as a function of all the relevant characteristics of the principal and agent:

$$y = \beta^p p + \beta^a a + \epsilon \tag{2.1}$$

p is a $K_p \times 1$ vector of principal or job characteristics. Examples might be the inherent riskiness of the principal's crop or franchise, the monitoring ability of the principal, the principal's cost of measurement or transactions, or the risk aversion of the principal.[8] Analogously, a is a $K_a \times 1$ vector of agent characteristics,e.g. the agent's risk aversion, productivity, or opportunity cost of effort. y represents the optimal contract given these characteristics. β^p and β^a are respectively K_p and K_a vectors of parameters that measure how principal and agent characteristics affect optimal contract choice. Throughout, we will consider ϵ to be measurement error or optimization error in y, uncorrelated with a and p. This loses little generality since p and a contain all characteristics relevant to contract choice and we will allow the possibility that elements of p or a are partially observed or unobserved by the econometrician.

In empirical work, transfer payments or franchise fees are often unobserved or ignored, so typically y might measure the share of output or revenue paid to the agent (e.g. royalty rate) or the discrete type of contract used (e.g. wage, share, or fixed-rent). As such, in actual application the contract choice equation may be non-linear. In this econometric section, we focus on the linear case (or an assumed linear approximation) because this

[8] We will include job or task characteristics as part of the principal's characteristics. This could become problematic if principals are matched with (or endogenously choose) particular tasks to contract out. Methods of dealing with this endogeneity problem would be similar to our methods for endogenous matching of principals and agents.

keeps identification issues as transparent as possible. In the empirical section, we also consider more appropriate non-linear models.

Note that we consider (2.1) a structural equation. In other words, p and a are the *fundamental* economic characteristics of principals and agents that affect contract choice. Therefore, one should think of the dimensions of p and a as being somewhat small. Even in the most complex theoretical literature, there is typically a relatively small number of economic fundamentals that determine the optimal contract. Clearly, if p and a are completely observable to the econometrician, (2.1) can be estimated consistently by either ordinary least squares or standard non-linear methods if y is latent or a non-linear function of the a's and p's. We now turn to the case where all these fundamental characteristics are *not* perfectly observed by the econometrician. Instead, we observe proxies for them. While the dimensions of p and a may be small, we may observe many such proxies.

3. The One-Sided Proxy Problem

We start with a one-sided proxy problem. In other words, while we do not perfectly observe some or all of the elements of a (or alternatively p), we do perfectly observe all elements of p (a). What we observe instead of a is a set of variables w that are potential proxies for the characteristics in a. We assume:

$$a = \Phi w + \eta \tag{3.1}$$

where w is a K_w vector of observed proxy variables, Φ is a K_a by K_w matrix of proxy coefficients, and η is a K_a-vector of proxy errors which we will assume is distributed mean independently of w with variance matrix Σ^{η} and mean 0. Importantly, and as standard when using proxies, the assumption that η is mean independent of w is essentially without loss of generality. The proxy equation is simply expressing the expectation of a given w. The only assumption involved is that this expectation is linear in w.[9]

Note that this general formulation allows (i) there to be more than one proxy variable for each characteristic a (the jth row of Φ contains more than 1 non-zero element), (ii) a particular proxy variable w_m to be a proxy for more than one of the characteristics a (the mth column of Φ contains more than one non-zero element), (iii) the possibility that there is no proxy for a particular a (the jth row of Φ contains zeros), or (iv) the possibility that we actually observe the true variable a (the variance of η_j is zero).

[9] In other words, it does not matter, e.g. if w is a "pure" proxy for a (i.e. $a = \theta w + \epsilon$, where w and ϵ are uncorrelated), whether w is a "measured with error" (i.e. $w = a + \epsilon$, where a and ϵ are uncorrelated), or in-between cases where ϵ is correlated with both a and w. Regardless, we can always write $a = E[a|w] + \eta$ where η is uncorrelated with w. As such, the only assumption we are making in equation 3.1 is that $E[a|w]$ is linear in w. On the other hand, the proxy-measurement error distinction and the causality issues associated with it do have important implications for the set of thought experiments that can meaningfully be considered with the model (e.g. "what would happen if we exogenously changed w?").

Substituting (3.1) into (2.1) gives

$$y = \beta^p p + \pi w + \beta^a \eta + \epsilon$$

where

$$\pi = \beta^a \Phi$$

is an M-vector of coefficients on the proxy variables. If p and w are uncorrelated with the resulting unobservable term $(\beta_a \eta + \epsilon)$, one can proceed to consistently estimate β^p and π using OLS or non-linear methods.

Given a-priori exclusion (i.e. $= 0$) restrictions on the proxy parameter matrix Φ (e.g. that a proxy variable proxies for only one particular characteristic) one can use the resultant estimates of π to potentially reject the null hypothesis that particular elements of β^a are zero. A sign restriction on the proxy relationship (i.e. on elements of Φ) can allow one to sign β^a. This is how most prior empirical work on contractual choice has proceeded in assessing whether particular characteristics are important determinants of contract form.

We now question the validity of this approach. By maintained assumption, both w and p are uncorrelated with the measurement/optimization error ϵ, and w is uncorrelated with the proxy error η. This leaves us with the question of whether p is uncorrelated with η. As η is part of a, this will generally *not* be the case if the variables p and a are correlated, i.e. if an agent's characteristics are correlated with the characteristics of the principal he is contracting with. If principals and agents were somehow randomly matched in a micro-economy, p and a might be uncorrelated. In actuality, there are likely benefits and costs of particular agents being matched with particular principals. In general, this should result in market outcomes where an agent's characteristics are correlated with those of the principal he is contracting with.

Incentives for Matching To exemplify the costs and benefits of matching, we briefly consider the Holmstrom and Milgrom (1987) moral hazard model in an economy with multiple principals and agents. Our goal here is not to build a structural model of matching and contracting to take to the data. It is simply to illustrate that there may be incentives for principals and agents to match in particular ways.

For simplicity assume that the costs of effort $c(\mu)$ are the same across agents and equal to $.5\mu^2$, and that the expected marginal benefit of effort is constant across jobs and equal to 1 (i.e. the expected revenue from output produced by effort μ is equal to μ). What does differ across principals and agents is agent i's risk aversion r_i and the variance in output of the job principal j is contracting for, σ_j^2. In terms of the above model, we can think of a and p each containing one element, i.e. $a = r_i$ and $p = \sigma_j^2$. Holmstrom and Milgrom show that if an agent with risk aversion r_i is contracting with a principal with output variance σ_j^2, the optimal share of output with which to compensate the agent with is:

$$\alpha = \frac{1}{1 + r_i \sigma_j^2} = \Psi_{ij}$$

In addition, one can show that the agent's effort given this optimal contract is also:

$$\mu = \frac{1}{1 + r_i \sigma_j^2} = \Psi_{ij}$$

and that the total certainty equivalent resulting from the contract is:

$$CE_{ij} = \Psi_{ij} - .5\Psi_{ij}^2 - \frac{r_i \sigma_j^2}{2}\Psi_{ij}^2 = \Psi_{ij} - .5\Psi_{ij}^2 \left(1 + r_i \sigma_j^2\right) = .5\Psi_{ij}$$

Now consider an economy with two heterogenous agents and two heterogeneous principals. Suppose first that $r_1 = 2$, $r_2 = .1$, $\sigma_1^2 = 2$, and $\sigma_2^2 = .1$. Consider the 4 possible combinations of principals and agents and the resulting total certainty equivalents: $CE_{11} = 1/10$ (the total certainty equivalent if agent 1 contracts with principal 1), $CE_{22} = 1/2, CE_{12} = 5/12,$ and $CE_{21} = 5/12$. Note that overall certainty equivalent in the economy is higher ($CE_{21} + CE_{12}$ =50/60>36/60=$CE_{11} + CE_{22}$) if the more risk averse agent gets "matched" with the less risky job. As such, we might expect this to be the outcome in this economy (more precisely, the outcome where agent 1 and principal 1 get matched is not in the core of this economy[10]). Interestingly, in this model it is not always best for the more risk-averse agent to be matched with the least risky principal. Take $r_1 = 3$, $r_2 = .5$, $\sigma_1^2 = 3$, and $\sigma_2^2 = .5$. This implies $CE_{11} = 1/20, CE_{22} = 4/10, CE_{12} = 2/10,$ and $CE_{21} = 2/10$. In this case, it is better for the reverse to happen ($CE_{21} + CE_{12}$ =8/20<9/20=$CE_{11} + CE_{22}$). The intuition in this second example is that there are severe enough risk sharing problems when *either* $r = 3$ or $\sigma^2 = 3$, such that little effort incentives can be given. As a result it is better for a social planner to match $r = .5$ and $\sigma^2 = .5$ and get strong incentives in at least one of the pairs.

This very simple example illustrates economic incentives that might lead to certain types of principals being matched with certain types of agents. In general, there is no reason to think that there will not be such incentives. If certain types of principals relatively prefer certain types of contracts and certain types of agents also prefer those types of contracts, one might expect it to be optimal for these types to be matched in equilibrium.[11] Though we will not delve significantly into how this matching might occur,

[10] Note that our preferred outcome is in the core. Transfer payments (i.e. a division of CE) that support this core allocation are: agent 1 - 10/60, agent 2 - 15/60, principal 1 - 10/60, principal 2 - 15/60.

[11] As another example, suppose agents differ in risk aversion but principals differ in monitoring ability (e.g. a fixed cost of measuring output when the contract is less than full share). One might expect it optimal for risk averse agents to match up with the principals with high monitoring ability, as both

we briefly mention some possibilities, as the theoretical matching literature has studied similar problems.

Extending on Shapley and Shubik (1972), Crawford and Knoer (1981) examine matching where both demanders and suppliers have idiosyncratic preferences for whom they get matched with. They show that this game has a non-empty core. What this means in our scenario is that there is a set of principal-agent matches such that no principal and agent can break out of their respective matches, negotiate with each other and do better. One to one matching with perfect information and a H-M moral hazard model of ex-post (after matching) contracting appears to fit into this framework.[12] Recent theory papers that much more fully address matching with general incentive problems include Shimer and Smith (1996) and Legros and Newman (1998).

A second possibility is a sophisticated hedonic equilibrium where there are prices (i.e. transfer payments) and an optimal share associated with every combination of principal and agent characteristics. Such an equilibrium should be in the core. Note that less sophisticated hedonic price equilibrium, without prices for every possible principal-agent combination, will not necessarily exist or result in core allocations.[13]

Another possibility, possibly most realistic in our opinion, is some type of search equilibrium where agents, for some search cost, get exogenously paired, observe each others characteristics, and decide whether to contract or search more depending on the "payoff" from the current pairing compared to expectations of that from more search. Again, we will not take a stand on how precisely such matching will occur. Our main points are that: 1) there are clearly incentives for certain types of agents to contract with certain types of principals, and 2) as such, it would be presumptuous to assume that the characteristics of an agent are uncorrelated with the characteristics of the principal they are contracting with.

---

relatively prefer sharecropping to full share contracts. Other examples are agents with lower costs of effort (or higher productivities) matching with principals with high measuring costs (both relatively prefer higher or full share contracts) or agents with credit constraints matching with principals with low measuring costs.

[12] The H-M model is very facilitating to this type of analysis because of its property that the optimal share, and thus CE, does not depend on transfer payments or reservation utilities, which would likely be determined endogenously in equilibrium.

[13] For example, consider a simpler hedonic price function that simply attaches a price (transfer payment to principal) to every share (share of output to be given to agent). We want to show that this simple price system may not result in a core allocation. Consider the H-M model with 3 principals ($\sigma^2 = .5, .5, 1$) and 3 agents ($r = .5, 1, 1$). The core in this case is where the 1's get matched with the .5's. Note that for both pairings ($\sigma^2 = .5, r = 1$), ($\sigma^2 = 1, r = .5$), the optimal share is the same (2/3) and the CE is the same (1/3). Under our simple price system, this implies that the transfer payment to the principal must be the same for both combinations. But if this is the case, a principal with $\sigma^2 = .5$ and the agent with $r = .5$ can necessarily break off and do better (at any possible price, in their respective supposed pairings, $CE_{\sigma^2=.5} + CE_{r=.5} = 1/3$. By breaking off, they can get a total $CE = 1/2$.)

Econometric Implications of Matching Returning to the econometric model,

$$y = \beta^p p + \pi w + (\beta^a \eta + \epsilon) \tag{3.2}$$

if there is such "matching" whereby agent characteristics a are correlated with principal characteristics p, then the unobserved component of a, i.e. the proxy error η, will generally be correlated with p.[14] This correlation between p and the econometric unobservable will bias estimates of β^p. In addition, as the matching process implies p is also correlated with w, the observed component of a, bias will also generally be imparted onto π. As the example above suggests, it will not always be clear who will be matched with whom. Therefore, theory can not be relied on to tell us which way these biases may go. This is a question that must be addressed empirically.[15]

Given that standard estimates of (3.2) will be biased, we turn to the question of how to obtain consistent estimates of β^p and π. We start with our matching equation—an equation describing the relationship between p and a. In a very reduced form approach to what is potentially a complicated matching equilibrium, suppose the equilibrium outcome is such that principals and agents are matched according to the multivariate function:

$$p = m(a, \nu) \tag{3.3}$$

ν is what we call "matching error", which allows for less than perfect matching of principals and agents. This might be particularly relevant if there is some search aspect of the matching process. It also permits there to be other variables that affect the matching of principals and agents but do not affect the contract choice equation. Examples might be idiosyncratic preferences to work in a certain location, job, or for a certain principal. The existence of ν, i.e. non-perfect correlation between p and a , is important empirically because without it, p and a would be essentially collinear, making any attempt at separating their individual effects on contract choice fruitless.

[14] Note that our discussion of the matching process and contracting has been under the assumption of perfect information, i.e. a and p are observable to all the actors in the economy. We wonder about the implications of parts of a and p being unobservable, particularly on the matching process. Perhaps the *best case scenario* for the validity of OLS estimates of (3.2) would be if the actors in the economy observe exactly what we as econometricians do (i.e. principals observe w of potential agents, but not a (i.e. η)). The question is then whether this unobserved η will affect whom an agent gets matched with. If η doesn't affect matching, it is not correlated with p and estimates of 3.2 are consistent. However, it is not clear that just because η is not observable, it will not affect matching. There is the potential for signalling type equilibria where η ends up being known in equilibrium (see, e.g. Ghosh, 1997).

[15] In some cases it is possible to make at least partial inferences. For example in the case where p, a, and w are one dimensional, we know that the OLS estimates of π are biased towards zero (Pf. Conditional on p, w is negatively correlated with η (positive if Φ is negative). Thus if 1) β^a and Φ are both positive (or both β^a and Φ are negative) , w and $\beta^a\eta$ are negatively correlated and the positive π is biased downwards, or if 2) β^a is positive and Φ is negative (or the reverse), w and $\beta^a\eta$ are positively correlated and the negative π is biased upwards QED). As a result, standard significance tests will not overreject a 0 null. On the other hand, the bias of β^p is not known (it depends on the direction of matching). Examining covariances can determine the sign of this bias, however (see below).

The *worst case* scenario for identification is if we assume that given an agent with characteristics a (and proxy w), the expected characteristics of the principal they will be matched with is linear in a, i.e. γa. In this case (3.3) is linear in a and ν where ν is uncorrelated with a, i.e.

$$p = \gamma a + \nu \tag{3.4}$$

Note that the interpretation of the data generating process here is that 1) agents are drawn from some population distribution of a, 2) a draw of the matching error ν determines the principal p they are matched with, and 3) a draw of ϵ determines contract choice.[16] In this case, combining (2.1), (3.1), and (3.4), results in the two reduced form equations:

$$\begin{aligned} p &= \theta w + \mu_p \\ y &= (\beta^p \theta + \pi) w + \mu_y \end{aligned} \tag{3.5}$$

where

$$\begin{aligned} \mu_p &= \gamma \eta + \nu \\ \mu_y &= (\beta^p \gamma + \beta^a)\eta + \beta^p \nu + \epsilon \end{aligned}$$

are unobservables and $\theta = \gamma \Phi$ is an M vector of reduced form coefficients. As μ_p and μ_y are independent of w, one can consistently estimate the reduced form coefficients θ and $(\beta^p \theta + \pi)$. However, without any instruments or covariance restrictions on the unobservables ν, η, and ϵ, one cannot separately identify the parameters of interest β^p and π. We examine possible covariance restrictions later. First we look at our preferred solution, that of instrumental variables.

Instrumental Variable Approaches What we need to separately identify β^p and π are instruments z that enter the matching equation but do not enter the contract choice or proxy equations. These instruments also need to be uncorrelated with both the proxy error (η) and the contract choice measurement/optimization error (ϵ). If we have a set of at least K (the dimension of p) such instruments, one can consistently estimate the contract choice equation (i.e. estimate β^p and π) by instrumenting for p. Note that this is the IV estimator generated by the orthogonality condition:

$$E\left[\beta^a \eta + \epsilon \mid w, z\right] = E\left[y - \beta^p p + \pi w \mid w, z\right] = 0$$

[16] Alternatively, one could consider a non-linear matching equation or a matching equation of the form $a = f(p, \nu)$ where ν is uncorrelated with p. In either of these cases, $E[p|a]$ will not generally be linear in a. These non-linearities can only "help" in estimation (as the nonlinearities might serve as instruments), but identification based on these non-linearities is clearly somewhat of a leap of faith and not robust to alternate functional forms of the contract choice equation.

What are potential instruments? We consider a couple of possibilities.[17] The first is that the matching equation is not linear in a. If the matching equation has higher order terms than does the contract equation, then these higher order terms can serve as instruments for p in the contract choice equation. As assuming such functional form restrictions can be tenuous, we prefer a second possibility—that one may observe other variables that affect the matching process. For example, one might use characteristics of a tenant that appear to affect matching but can reasonably be excluded from entering the contract choice equation.

Perhaps most interesting is the case where observations come from different geographical or temporal markets. If the population distribution of principal or agent characteristics differ across these markets (e.g. the distribution of types (riskiness) of crops), it is natural that the matching equation should differ across the markets. For example, consider two geographically distinct markets—both have the same population distribution of risk aversion across agents, but in the second market the distribution of riskiness of tasks (i.e. principal characteristics) has a higher mean. For a given level of risk aversion, a tenant in the second market will tend to get matched with a more risky task.

While different means of these characteristic distributions generate differences in the intercept term of matching equations, different variances of these distributions tend to generate different slope coefficients. If, for example, all jobs in the second market have the same level of riskiness, the slope coefficient in the second market will be zero. In the above example, differences in the distribution of p causes differences in the matching equation across markets. Even if the distribution of p is very similar across markets, differences in the distributions of w will also tend to generate differences in the matching equations.

If the slopes and/or the constant terms of the matching equations vary across markets,

[17]Note the similarities between potential instruments here and instruments that have been proposed in the hedonic literature (e.g. Epple 1987; Bartik, 1987; Kahn and Lang, 1988). Interestingly, one could potentially analyze contract choice data in a hedonic framework. While such a method would address matching more structurally than does ours, it would also have a number of limitations. For example, (i) it would be much more data intensive, as data on hedonic "prices" (e.g. transfer payments) of all possible contracts would be necessary. Usually data on shares or royalties is much easier to come by (for example, our dataset contains no information on transfer payments), (ii) it relies on more assumptions about matching - with a contract choice equation the only assumption is that an optimal contract is chosen *conditional* on a match, (iii) it does not directly estimate the optimal contract choice equation (which are often the parameters of interest), and (iv) it would be considerably more complicated, since as noted in section 3, one generally needs to estimate a hedonic price function over *all* principal and agent characteristics. Brandt and Hosios (1996) is an interesting paper that uses this hedonic approach. In their case, the complexity of the hedonic approach is mitigated because in their particular model, conditional on the contract terms (price and payoff structure), principals and agents do not care who they are matched with. As such, they can use a hedonic only over contract terms. This is not true in our model with heterogeneity in riskiness and risk aversion. For example, conditional on a transfer payment and share, a principal *does* care what type of tenant he/she is matched with. As a result, we would need to estimate a hedonic price function over all principal and tenant characteristics.

we have two sets of potential instruments—1) market dummies, and 2) market dummies interacted with w.[18] The next question is whether these instruments can reasonably be excluded from the other equations in the system. First, consider the structural contract choice equation. Recall that the contract choice equation specifies optimal contractual form as a function of *all* the *fundamental* economic characteristics of the principal and agent. As there is no compelling argument to think nor model to suggest that market of residence is one of these fundamental characteristics, these market variables can reasonably be excluded virtually by definition.

This, unfortunately, is a bit misleading, as it is conceivable that some of these fundamental characteristics (elements of p and a) are "endogenous" to the market, i.e. they are actually determined by the market the actors are in. We can think of one particular example. In a matching equilibrium, it is likely that reservation utilities of principals and agents are endogenously determined. These different reservation utilities may be associated with different transfer payments to the "same" agents across different markets. For example, equilibrium transfer payments to "valuable" agents (e.g. less risk-averse ones) who are relatively scarce in one market may be higher than in another market where they are less scarce. Equilibrium transfer payments varying across markets is not a direct problem (at least when the variable y is the optimal share of output[19]). On the other hand, a problem would arise if differences in transfer payments indirectly affected the optimal contract choice y. The primary example we can think of is through income effects—i.e. suppose transfer payments affect "net" tenant wealth, which in turn affect tenant risk aversion, in turn affecting optimal contract choice. On the other hand, many models do not have this effect, e.g. in the Holmstrom and Milgrom moral hazard model the agent's risk aversion is not affected by the transfer payment. This should also be the case in other CARA (constant absolute risk aversion) models. Even in more general models, we feel that these should be small, indirect effects that will be unlikely to cause significant problems.

---

[18] One can use the same market variable dummies and interactions with w's to instrument for all p's. Thus, as the dimension of p increases, one needs either the number of non-market based instruments to increase, the dimension of a and w to increase, or the number of markets to increase for continued identification. On a similar note, our methodology is also applicable if the contract choice equation includes interactions between pricipal and agent characteristics, e.g.

$$y = \beta^p p + \beta^a a + \beta^{pa}(p \quad a) + \epsilon$$

$$= \beta^p p + \pi w + \beta^{pa}\gamma(p \quad w) + (\beta^{pa}(p \quad \eta) + \beta^a \eta + \epsilon)$$

As the interaction terms are also endogenous to the error term, one needs additional instruments - natural such instruments are the instruments for p interacted with w. Note that in this case: (i) some of the market based interaction instruments can become redundant, and (ii) this become problematic if the proxy error η is heteroskedastic (w.r.t. w), as the expectation of the error term includes a σ_η^2 term.

[19] Note that we probably wouldn't advocate use of these market based instruments if y were, for example, the transfer payment associated with the optimal contract.

Secondly, we need the assumption that these market based instruments do not enter the proxy equation. What this requires is that the mean of the proxy equation or the slope of the proxy equation do not vary across markets. For example, the expected value of the unobserved component of risk aversion (η) cannot vary across markets.[20] This suggests we want the markets as isolated as possible. We would not want, for example, tenants to be moving around based on economic considerations; as a result particular types of tenants might want to move to markets where payoffs to those characteristics are high. While these assumptions may not always hold, they are essentially on the same level as those in a standard proxy problem, i.e. that the proxy errors are uncorrelated with other assumed *exogenous* variables in the model. Essentially, while we are allowing an agent to match within a market endogenously, we are assuming that the market the agent is in is exogenous. These assumptions are also analogous to the "market-instrument" assumption often used to solve the hedonic identification problem (see Kahn and Lang 1988; Brandt and Hosios 1996). In that literature, distributions of unobserved consumer and producer preferences or characteristics are assumed identical across markets.[21] Lastly, note that if one has a large number of markets it may be safer to use characteristics (e.g. 1st and 2nd moments) of the distributions of p as instruments instead of the market dummies. As with the market dummies, one can use both moments and moments interacted with w. This is safer because it relies on less restrictive assumptions. For example, the mean of η *can* vary across markets as long as this mean is uncorrelated with the moments of the p distribution[22].

Covariance Restrictions Approach What is attractive about the IV solution is that it does not depend on the functional form of the matching equation or on covariance restrictions on the structural errors. However, if one does not have or trust any of the afore-mentioned instruments, it is possible that β^p and π can be identified by covariance restrictions. As an example, assume the linear matching equation (3.4), suppose a, w, and p are scalars (one can put a constant term in a), and make the covariance restriction that ν, η, and ϵ are uncorrelated and that σ_ϵ^2 is known (e.g. ϵ is identically 0 (i.e. there is no measurement or optimization error in the contract choice equation)). Note that this assumption requires there to be no other unobservables that affect both matching and the

[20]There is actually not a problem if the variance of the proxy error differs across markets. As such, the differences in the distributions of principal and agent characteristics across markets we need to have (in order to actually have potential instruments) must be "caused" by only observables (e.g. w and p) or differences in the variances of unobservables (the proxy errors).

[21]This assumption would similarly be invalid if consumers, based on unobservables, choose what market to be in (e.g. in order to get a good price on the product they want).

[22]As such, this approach still wouldn't solve problems arising from economic based migration. The reason is that migration (the mean of η) would likely be correlated with the distribution of land. Note that with a small number of regions, using these market moments will be identical to using market dummies (if the number of moments equals the number of markets minus 1).

contract choice equation. In this case the covariances of the reduced form equations (3.5) will separately identify β^p and π (See Appendix). Essentially the identification comes from the fact that the observed correlation between p and w tells us exactly what the correlation between p and η is. Even if one is unwilling to assume σ_ϵ^2 is known (assuming ϵ, ν, and η are uncorrelated), one can sign the biases on β^p and π (see footnote 15 above).[23]

These covariance restrictions bring up an important point regarding the "first-stage" regression in the IV solution. As usual, examination of this regression (in our case simply regressing principal characteristics on observed agent characteristics, market dummies, and market dummies interacted with agent characteristics) is of interest to determining whether one has significant instruments. But suppose one finds no significant correlation between observed agent and principal characteristics. While this may rule out some potential instruments, it also suggests that there may be no matching. Under the above covariance restrictions[24], it implies that $\gamma = 0$ and thus there is no endogeneity problem in the contract choice equation. Clearly, a first step with this type of data is to examine correlations between principal and agent characteristics.

4. The Two Sided Proxy Problem

This section briefly discusses the case where characteristics of *both* principals and agents are not perfectly observed. This considerably complicates the problem, as if there is matching, essentially all variables are endogenous. As before, observed principal characteristics are correlated with unobserved (or unobserved portions of) agent characteristics. In addition, observed agent characteristics are correlated with unobserved principal characteristics. To exemplify, suppose

$$p = \quad m + \tau \tag{4.1}$$

where p are the true principal characteristics, m is a K_m vector of proxies for these characteristics, and τ is a K_p vector of principal proxy errors that are independent of m. Our linear matching equation:

$$p = \gamma a + \nu$$

now becomes

$$m + \tau = \gamma(\Phi w + \eta) + \nu$$

As before, m is correlated with the agent proxy errors η, and in addition, w is correlated with the principal proxy errors τ. Since the contract choice equation is now:

$$y \quad = \quad \beta^p p + \pi w + \beta^a \eta + \epsilon$$

[23] One can increase the dimensionality of p and keep this identification (assuming the covariance matrix of the error term in the now multivariate matching equation is diagonal), but increases in the dimensionality of a are problematic, even if the covariances of the multiple proxy errors are assumed zero.

[24] Plus the assumption that one has at least one valid (significant) proxy for each agent characteristic. Otherwise, it might be that the proxies are simply not picking up the matching.

$$= \psi m + \pi w + (\beta^a \eta + \beta^p \tau + \epsilon)$$

where $\psi = \beta^p$, both m and w are correlated with the resulting error term and nothing on the right hand side is econometrically exogenous.

If one has enough markets and there is enough variation in matching equations across markets, one can also use market based instruments for the two-sided case, though the demands of the instruments are clearly greater. A second solution pertains if we have multiple observations on either the principal or the agent and we are willing to assume that the proxy error for that actor is constant over these observations. One example might be a principal contracting over multiple pieces of land with different tenants. If τ represents an unobserved component of a principal which is fixed over the multiple pieces of land, we can difference it out of both the matching and contract choice equations, solving one of the endogeneity problems. As long as instruments are varying across tenants for a given landlord, one can use IV to solve the remaining problem. Note that this rules out market dummy instruments, but does allow market slope instruments. Clearly, one could use time series data in a similar way if one is willing to assume that τ is constant for a given landlord over time. As usual, this differencing methodology is only applicable for linear (or transformed to linear) models[25]. In non-linear models, one will need to move to random effects formulations.

5. Data

5.1. The Sample

This paper examines matching and its implications on contract choice equations using an interesting historical dataset on agrarian contracts—the Florentine *Catasto* of 1427, a comprehensive census and property survey of Tuscany. The sample consists of landlords living in the towns of Florence, Pescia, and San Gimignano.[26] Landlords often contracted with more than one tenant—the data consists of 902 land plots/contracts owned by 128 landlords.[27] The countrysides of these towns were isolated from each other economically

[25] Note also that use of fixed-effect logit models will typically not be applicable unless all the *observed* (as well as unobserved) principal characteristics are invariant over observations. Unless this is the case, one still has an endogenous m, which cannot be dealt with in the conditional logit model.

[26] Pescia and San Gimignano were two towns under Florence's rule. The population of Pescia was 532, of San Gimignano 576, and of Florence itself, more than 9000. For the survey, each head of household had to declare: the houses, lands, and draft animals he owned; the crops grown and the agrarian contracts used; the average crop yields over the previous three years (though unfortunately not the crop yield in each of the 3 years); his debts and credits; his shares of commercial partnerships; and his profession. Furthermore, he had to report the composition of his family by name, age, sex, and relationship to himself. Sixty thousands households were surveyed in the *Catasto*.

[27] In order to ease data collection, we oversampled from the more wealthy landlords. These are exactly the kind of landlords that owned many land holdings and leased their land plots out to share croppers

and migration rates were fairly low: 3 percent of peasant households living in the countryside of Florence declared to have emigrated from other Tuscan towns or from other places. The corresponding percentages for San Gimignano and Pescia were 6 and 13 percent, respectively. In light of the discussion in Section 3, this separation is important as it supports towns as potential instruments in estimation.

The sample contains both share and fixed-rent contracts. Although the data does not contain information on exact shares, most were 50/50.[28] The crop type on a particular land plot was classified as either "vines," "cereals," or "mixed," depending on whether vines (or other perennial crops), cereals (or other annual crops), or both types of crops were cultivated.[29]

The distinction between cereals and vines is relevant both for the risk sharing hypothesis and the multitasking model of Holmstrom and Milgrom. First, while cereals were also subject to weather variability, vines were apparently much more sensitive.[30] As such, all else equal we might expect vines to be more conducive to share rather than fixed rent contracts. Second, vines had interesting multitasking aspects. Peasant tenants could boost current production by pruning in a certain way or putting manure near the roots. However, the vines would have been damaged and would have produced less output in the successive years (Galassi 1998). There was almost a trade-off for peasant tenants between either putting their effort in maximizing current production or in maintaining and improving the assets for future production. Thus, viticulture shares many characteristics of the multitasking setup of the Holmstrom and Milgrom model, and landlords of vineland might be hesitant to sign contracts with strong incentives for current production (i.e. fixed rent contracts).[31]

Both the potential multitasking and risk-sharing effects of vines suggest similar outcomes—less fixed-rent contracts on vines. Therefore, considering the crop type alone is not a conclusive test of the risk sharing hypothesis. Thus, we also consider the tenant's wealth as a potential proxy for his risk aversion. If risk-aversion is an important determinant of

and fixed rent tenants. We set our sampling criterion (i.e. a lower bound on landlord wealth) such that approximately 85% of share or rented plots in each town were potentially in the sample. Our hope is that this large percentage should imply small potential selection problems.

[28] As often the case with data on contracts, the preponderance of a few particular contracts (in our case fixed-rent and half share) does not completely coincide with theory models that would suggest we should observe a wide range of shares (or non-linear contracts). One argument is to appeal to institutional restrictions that limit "optimal" choice to the optimum of a small set of commonly used contracts.

[29] Typically the nature of the land on a farm determined the crops to be planted, so it is reasonable to think of the type of crop to be exogenous to the land and landowner. However, since we end up modelling crop type as endogenous anyway (due to endogenous matching), this assumption is not completely necessary.

[30] See Galassi (1998). Higher output variability doesn't necessarily transform into lower *revenue* variability, but unfortunately we do not have evidence on this.

[31] Though we do not have data over time, there is evidence that vines were more likely to have contracts longer than one year (Bandiera 1998). This evidence would also seem supportive of multitasking.

contract form and wealth is a sufficient proxy for risk-aversion, wealthier tenants should be more likely to engage in fixed-rent contracts.[32]

Lastly, the data includes two variables that might shed light on potential landlord monitoring or measurement problems. A landlord's occupation and the number of adult children in the household are potential proxies for this monitoring ability. It might be more costly for landlords practicing nonagricultural occupations (such as merchants, notaries, medical doctors, or artisans) to monitor tenants, and less costly for landlords with more adult children to help monitor the tenants. Higher costs of monitoring would presumably lead to higher likelihoods of fixed rent contracts.

5.2. Summary Statistics

Chart 1 presents summary statistics of all variables used. The 4 cross-tabulations in Table 1 (for the aggregate and for the 3 towns individually) provide some interesting information on correlations in the sample. Three interesting correlations emerge, holding both in the aggregate and at the town level. First, land plots with vines appear mostly associated with share contracts, while land plots with cereals were most often leased out under fixed-rent contracts. Out of 902 observations, 482 land plots with vines or mixed crops were share cropped, and 28 were leased out to fixed-rent tenants. As for land plots with cereals, 83 were share cropped and 309 were leased out to fixed-rent tenants. Second, looking across contracts for a given crop type, mean tenant wealth is typically higher under fixed-rent contracts than share contracts. Third, Table 1 indicates correlation between a tenant's wealth and the type of crop they are cultivating. Poorer tenants primarily cultivated land plots with vines only or farms with vines and cereals. Wealthier tenants mainly cultivated plots with cereals only. This last correlation is suggestive of matching between landlords and tenants: tenants with certain characteristics appear to be matched with specific crops.

Lastly, note that the distribution of land types is very different across towns. Pescia had about 75% pure cereal plots, had very few mixed plots—the rest (20%) were plots with vines. San Gimignano had almost 85% mixed plots, 10% pure cereals, and 5% pure vines. Florence was somewhat in between, with 55% mixed, 38% cereals, and 8% vines. These differences in land distributions can give exogenous variation in crop-type that can help control for endogenous matching.

6. Empirical Results

We present two sets of empirical results. The first are linear models similar to those discussed above. While there is a considerable amount of discreteness in our data (type of

[32] Wealth has been used as a proxy for risk-aversion in a number of previous empirical studies, e.g. Laffont and Matoussi (1995).

land, type of contract), the linear models make identification issues clear and transparent. We then turn to non-linear models that are more appropriate given our data. Since sources of identification are more obscured in these non-linear models, we devote a section to addressing the issue.

6.1. Linear Models

Naive Contract Choice Equations Our first econometric take on the data is to estimate simple contract choice equations, ignoring the possibility of matching and subsequent endogeneity problems. Column 1 of Table 2 presents results of a linear probability regression of contract choice (0=share, 1=fixed rent) on town dummies, tenant *wealth*, and a crop type variable *vines* (=0 if cereals, =0.5 if mixed, =1 if vines). Our results are robust to alternative definitions of the *vines* variable.[33,34] The results confirm the casual evidence of Table 1—moving from cereals to mixed crops to vines appears to decrease the probability of fixed-rent contracts, while increases in tenant wealth increase the likelihood of fixed rent contracts. The negative coefficient on *vines* might suggest either (i) vines are more risky and there are thus more incentives to risk share on them, or (ii) support a multitasking argument through which landlords are hesitant to use incentive laden contracts on vines. If *wealth* is a good proxy for risk aversion, the positive coefficient would seem to lend some support to the risk sharing hypothesis. Foreshadowing our second set of estimates, Column 3 repeats the analysis with a more appealing probit model of contract choice—again *vines* is very significant.

Columns 2 and 4 run the same models using only the data for Pescia and San Gimignano. There are a couple of reasons for doing this. First, Florence was considerably different from the other two towns. Not only was it much bigger both in size and population, but it was also the commercial center of Tuscany. Second, as evidenced by the estimates, there do appear to be differences, particularly in the coefficient on *wealth*. This coefficient is considerably smaller than that in the complete regression, *not* statistically significant in the probit model, and marginally significant in the linear probability model.[35]

---

[33] For example, when we set vines = 0, 0.25, 0.75 or 1 when there are mixed crops. Note that when we set vines=1 for mixed crops, this corresponds with the vines =1 when there are any vines at all on the plot. The reason we make a-priori restrictions on this variable and do not estimate coefficients on vines and mixed crops separately is because it adds an extra endogenous variable to the system later.

[34] Since these are linear probability models with resulting heteroskedasticity, all reported standard errors and p-values are bootstrapped.

[35] The differences between the full-sample results and the Pescia-San Gimignano become important later when we consider the potential IV identifying restriction that the wealth coefficient is equal across towns. See column 7 for a regression where the wealth coefficient is allowed to vary across all 3 towns. While the Florence coefficient appears considerably different (and is significantly different than the other two), the difference between the Pescia and San Gimignano coefficients is not significant. Note also that we include region dummies in all the contract choice equation. The reason we relax this potential

Thus, while the effect of *vines* on contract choice seems very strong regardless of specification, support for the risk-sharing hypothesis through the *wealth* variable is definitely mixed, particularly in Pescia and San Gimignano. Columns 5 and 6 obtain similar results with linear fixed effect models. These specifications allow for unobserved variables that are constant for a given landlord (e.g. differences in monitoring or transactions costs) and might be correlated with *vines* or *wealth.*

Matching Equations We now consider the possibility of endogenous principal-agent matching. In our case, particular concern arises from the realization that tenant wealth is probably not a perfect proxy for tenant risk aversion. If tenants match with landlords based on an unobserved component of risk aversion, the coefficients in Table 2 on both *vines* and *wealth* will be biased. For the moment, consider the one-sided proxy problem—i.e. assume that landlords have only one relevant characteristic, land type, and this land type is perfectly observed by us as econometricians.

As suggested in section 3, a first step towards assessing potential matching problems is to examine correlations between observable principal and agent characteristics. Column 1 of Table 3 addresses this question by regressing *vines* on town dummies and tenant *wealth.* The results confirm the correlations in Table 1 in that they show a very strong, significant, negative relation between the two, i.e. it is the less wealthy tenants that appear to end up on vines.

At the very least, this regression suggests that there *is* matching between principals and agents. It does not tell us why there is such matching, although with more heroic assumptions, one can infer more from this negative relationship. First, if risk effects were very important and vines were considerably more risky than cereals, one might expect the *reverse* relationship, i.e. less risk averse tenants ending up on more risky crops.[36] On the other hand, a strong multitasking story (with little or no difference in the riskiness of crops) might suggest that more risk averse tenants end up on vines. The reasoning is that, all else equal, both landlords owning vines, and very risk averse tenants, relatively prefer share contracts. These two arguments suggest that under the hypothesis that matching can *only* be caused by risk issues or multitasking, the negative relation in the data might be more supportive of multitasking. Of course, this argument needs to assume away any other potential reasons for matching, e.g. less wealthy tenants like wine, or that there are other economic reasons why less wealthy tenants would prefer to cultivate vines.[37]

identifying restriction is that it is strongly rejected by the data (both in OLS and in future IV regressions under overidentifying restrictions).

[36] Although some of the results of section 3 show that this need not be the case.

[37] It is important to note that while the above was a brief stab at inferring something from the direction of matching, the remainder of our current empirical work assumes *nothing* about the causes of matching. We are simply concerned with the fact that matching appears to exist and its implications for our estimates of the contract choice equation.

Covariance Restrictions Regardless of the actual cause of matching, the estimated direction of matching has implications for the contract choice equation estimates if one is willing to make structural covariance restrictions. Following the covariance assumptions of section 3, we might assume that the errors in the observed contract choice and matching equations are correlated, but only through an unobserved component of tenant risk aversion. Under this assumption, our coefficient on *wealth* will be biased towards zero. The negative correlation between *vines* and *wealth (*the observed component of risk aversion) indicates negative correlation between *vines* and the contract choice error term, resulting in a negatively biased *vines* coefficient. This at least casts doubt on the strongly negative *vines* coefficient, and also suggests that the effect of risk aversion may be stronger and/or more significant than it appears. If we make the stronger assumption that there is no erropr term in the structural contract choice equation (i.e. $\sigma_\epsilon^2 = 0$ - in other words that the only error in the *observed* contract choice equation is due to an unobserved component of risk aversion), the coefficient on *wealth* becomes positive and significant (0.08 with a bootstrapped p-value of .07) while the *vines* coefficient becomes very small and insignificant (-0.01 with a bootstrapped p-value of .974) .[38] While these are strong assumptions, it is interesting that these covariance restriction results are consistent with the IV results that follow.

Linear IV Models Given the apparent existence of matching and the resulting endogeneity of *vines*, section 3 suggests the possibility of instrumenting using cross-region differences in the matching equation. Our hope is that significant such differences will provide instruments for the *vines* variable. Recall that with this method there is no need for covariance restrictions and there is no need to make assumptions about the causes of matching.

Note the "natural experiment" we are looking for here. We would like to take a tenant with a certain level of wealth (and expected risk aversion) and move him from town to town. If there are differences in the regional distributions of principal and agent characteristics (and thus differences in the equilibrium matching equation), this movement may lead to the tenant ending up on different types of land. This is the source of exogenous variation that we want in order to identify the *vines* coefficient.

Column 2 of Table 3 assesses whether we have such differences by examining a less restrictive matching equation allowing for regional differences in both the intercept and slope coefficients. Though there is a significant difference in constant terms, our allowance of regional dummies in the contract choice equation means that these differences are not helpful as instruments. More important for our purposes are the differences in the matching function *slope* coefficients. γ_1^S , γ_1^P, and γ_1^F represent the slope term in San Gimignano, Pescia, and Florence, respectively. The estimates suggest that the relation

[38] This is for the San Gimignano-Pescia subsample. Results were similar when we ran these on the two towns (San Gimignano and Pescia) individually.

between *wealth* and *vines* is negative in all three, but strongest in San Gimignano, next strongest in Florence, and weakest in Pescia. The three terms are all significantly different from each other, suggesting the use of their differences as instruments – if they can be excluded from the contract choice equation.[39] Note the restrictions involved in making this exclusion. First, risk aversion's effect on contract choice must be the same across the relevant towns. Second, the proxy relationship between wealth and risk aversion needs to have the same slope across towns.[40] Recall that these restrictions are one of the reasons why we examine both the full sample and the restricted Pescia-San Gimignano sample. The Pescia-San Gimignano subsample results are based on the weaker restriction that only the coefficients in these two towns are the same.

Columns 3 and 4 contain linear IV results for the full and Pescia-San Gimignano samples, respectively. Columns 5 and 6 are IV fixed effects models. These IV fixed effects models allow for two-sided matching to the extent that principal unobservables are constant across different tenants. This could correspond with principals differing in monitoring or transactions costs. Such differences might affect both matching and the optimal contract choice. Comparing the IV results to their corresponding OLS results in Table 2 shows a number of differences. In all cases the *vines* coefficient becomes smaller and considerably less significant. Similarly, in all cases the *wealth* coefficient increases in magnitude. These changes are most apparent in the *least restrictive* specification, the fixed effect model with the Pescia-San Gimignano data (Column 6).[41] Comparing these estimates to their OLS analog (Table 2-Column 6) the estimated coefficient on *wealth* more than doubles (from 0.0242 to 0.0530) and moves from borderline significance (one sided p-value of 0.0447) to clear significance (p-value = 0.0048). Note that the differences between the OLS and IV results are indicative of negative correlation between the errors in the contract choice and matching equations. Given the signs of the γ_1's (-) and β_2 (+), this negative correlation is consistent with an unobserved component of risk aversion entering both these error terms.[42]

In summary, the linear models suggest that (i) there is matching, and (ii) controlling for this matching *does* make a difference. Particularly in the Pescia-San Gimignano dataset,

[39] Bootstrapped p-values for the significance of these differences are $\gamma_1^P - \gamma_1^S$—0.001, $\gamma_1^P - \gamma_1^F$—0.039, and $\gamma_1^S - \gamma_1^F$—0.0068.

[40] Given our inclusion of region dummy constants in the contract choice equation, it is not essential that the proxy relationship have the same mean across markets. Any differences will be picked up by these constants. It is also not necessary for the unobserved component of risk aversion to have the same variance across markets.

[41] This is the least restrictive specification since it does not assume that the contracting environment is the same in San Gimignano and Pescia as it is in Florence, and that it allows for unobserved landlord characteristics.

[42] Note that with the IV estimators, there is nothing forcing the correlation and resulting biases to go in this direction. Thus, the instruments are somewhat supportive of the conclusions made with covariance arguments. The estimated correlation coefficients between the error terms in the Pescia - San Gimignano models are -0.179 for the IV model and -0.434 for the IV Fixed effects model.

the OLS estimates overestimate the effect of *vines* effect and underestimate the effect of *wealth*, both in value and particularly in statistical significance. Again, both the IV and covariance restriction results support this conclusion.

6.2. Nonlinear Models

FIML Models The linear IV models ignore a considerable amount of discreteness in the model. The contract variable is one of two discrete types, while land type (*vines*) is of three discrete types. To control for endogenous matching accommodating these non-linearities, we need to consider the full system of equations. Assume that contract choice is given by the probit:

$$y_i = I(\quad \beta_0^k + \beta_1 v_i + \beta_2 w_i + \epsilon_i^1 > 0 \quad) \tag{6.1}$$

and the matching equation is given by the ordered probit:

$$v_i(w_i, k, \epsilon_i^2) = \begin{cases} 1 & \text{if } \gamma_1^k w_i + \epsilon_i^2 > \overline{C}^k \\ 0.5 & \text{if } \underline{C}^k < \gamma_1^k w_i + \epsilon_i^2 < \overline{C}^k \\ 0 & \text{if } \underline{C}^k > \gamma_1^k w_i + \epsilon_i^2 \end{cases} \tag{6.2}$$

where k indexes the 3 (or 2) different towns.

A few remarks are in order. First, both the slopes (γ_1^k) and cutoffs ($\overline{C}^k$ and $\underline{C}^k$) in the ordered probit matching equation are allowed to depend on town k. Because of the discrete nature of the processes, the variances of the assumed normal unobservables ϵ_i^1 and ϵ_i^2 need to be normalized to one, but we estimate a correlation coefficient between the two unobservables. This correlation is particularly important to accommodate because at the very least, ϵ_i^1 and ϵ_i^2 both include the unobserved component of tenant's risk aversion. Lastly, note that it is the "ordered" variable v_i that enters the contract choice equation, *not* the latent variable determining v_i. This corresponds with our interpretation that it is the discrete move from vines to cereals that has an effect on contract choice.

For these FIML models we focus attention on the Pescia-San Gimignano dataset, both because it relies on less identifying restrictions and because this is where controlling for matching makes the biggest difference.[43] Column 1 of Table 4 re-estimates the naive probit contract choice equation ignoring the endogeneity of *vines*—β_1 is very significant, but β_2 is not. Column 2 contains results of the ordered probit matching equation estimated independently. Note that both the upper and lower cutoffs and the coefficient on *wealth* in the ordered probit differ significantly across towns. This suggests that we do have potential instruments for v_i.

Column 3 estimates the full model, allowing for correlation between the error terms. The pattern is very similar to the linear models, if not a bit more dramatic. The estimate

[43] Results of the FIML models on the full dataset are available from the authors.

of β_1 more than halves and becomes insignificant, while the estimate of β_2 almost doubles and becomes significant. The correlation term is negative and significant, again consistent with an unobserved component of risk aversion. Again, while the naive estimates suggest that the effect of *vines* is the strongest statistical relation in the data, controlling for endogenous matching suggests that the *wealth* effect predominates.

Identification Caveats and Robustness The intuition behind the FIML estimates in Column 3 is the same as in the linear models. We want cross-region differences in the matching equation to identify the contract choice equation. Interestingly, there are additional, perhaps unwanted, sources of identification in the full model. These arise from the inherent non-linearities in the structural matching equation. Recall that in a model with a linear matching equation, no instruments, and no covariance restrictions, it is impossible to distinguish the direct effect of w_i on the contract variable $y_i^* = \beta_0^k + \beta_1 v_i + \beta_2 w_i + \epsilon_i^1$ versus the indirect effect of w_i through v_i. One ends up with a $(\beta_1\gamma_1 + \beta_2)$ term on w_i that cannot be disentangled. In contrast, in the above *non-linear* matching model, the direct effect of w_i on y_i^* is still assumed linear, but the indirect effect, i.e. $E[\beta_1 v_i \mid w_i]$ is non-linear through the ordered probit structure. Thus, the system of (6.1) and (6.2) will generally be "identified" even without standard instruments. However, we would not describe this as "good" identification, as it is not even robust to alternative functional forms for w_i in the contract choice equation. As such we want to be careful in assessing what is telling us what about β_1 and β_2.

One way to separate what the "good" instruments are telling us from what the "bad" ones are would be to allow w_i to directly enter the contract choice equation with arbitrary functional form, i.e. non-parametrically. In this case, any non-linear function of w_i implicitly enters the contract choice equation. Since we do not have all that much data, we take an alternative approach. We include w_i in the contract choice equation in a very particular non-linear way—specifically one that imitates w_i's non-linear effect through v_i. Consider the function:

$$OP(w_i, \epsilon_i^2) = \begin{cases} 1 & \text{if } E\left[\gamma_1^k\right] w_i + \epsilon_i^2 > E\overline{C}^k \\ 0.5 & \text{if } E\underline{C}^k < E\left[\gamma_1^k\right] w_i + \epsilon_i^2 < E\overline{C}^k \\ 0 & \text{if } E\underline{C}^k > E\left[\gamma_1^k\right] w_i + \epsilon_i^2 \end{cases}$$

where the E are expected values of the city dummy instruments (e.g. $E\gamma_1^k$ is the population weighted average of the 2 (or 3) γ_1^k's). Note that $OP(w_i, \epsilon_i^2)$ is a non-linear function of w that is very similar to $v_i(w_i, k, \epsilon_i^2)$ and *exactly equal* to it if the city coefficients are all the same. Thus, the contract choice equation:

$$y_i = I(\quad \beta_0^k + \beta_1 v_i(w_i, k, \epsilon_i^2) + \beta_2 w_i + \theta OP(w_i, \epsilon_i^2) + \epsilon_i^1 > 0 \quad)$$

will *not be identified* if the matching equation does not depend on region. Column 4 presents estimates of this model. The coefficient on vines is now significant and θ is positive and significant. Note that the overall effect of wealth on contract choice in this model is a combination of the linear term β_2, which is positive, and the non-linear term θ, which ends up being negative since wealth has a negative effect on $OP(w_i, \epsilon_i^2)$. This overall effect ends up positive and very significant (significance was computed by bootstrapping). We conclude that the insignificant coefficient on *vines* in Column 3 may be misleading through contamination with some non-robust sources of identification. When we allow for endogenous matching and utilize only good sources of identification, both *vines* and *wealth* appear statistically significant.

Random Effects Models Lastly, we consider allowing unobserved principal characteristics in the non-linear model. Because of the non-linearities and our short panel "length", it is not possible to do fixed effects. Therefore, we turn to a random effects formulation. One interpretation is that this random effect, i.e. α_j, is monitoring or measuring ability of the landlord that might increase the probability of share rather than fixed rent contracts. With matching, one might expect a landlord's α_j to be correlated with the risk aversion of the tenants they match with. For example, risk averse tenants might match with landlords with good monitoring ability, as both relatively prefer share contracts. This is the two sided proxy problem of section 4. Unobserved tenant characteristics are correlated with observed landlord characteristics and unobserved landlord characteristics are correlated with observed tenant characteristics. We deal with this in an admittedly ad-hoc way, simply including α_j on the right hand side of both (6.1) and (6.2) and modelling α_j as a linear function of the mean wealth of tenants contracting with that landlord (and other landlord variables) plus an unobserved component.[44]

Results of the random effects models are presented in Table 5. γ_3 is the coefficient on the random effect in the matching equation, σ_α is the standard deviation of the unobserved component of the random effect, and the coefficient on the random effect in the

[44] Formally we have:

$$v(w_i, k, \epsilon_i^2) = \begin{cases} 1 & \text{if } \gamma_1^k w_i + \epsilon_i^2 + \gamma_3 \alpha_j > \overline{C}^k \\ 0.5 & \text{if } \underline{C}^k < \gamma_1^k w_i + \epsilon_i^2 + \gamma_3 \alpha_j < \overline{C}^k \\ 0 & \text{if } \underline{C}^k > \gamma_1^k w_i + \epsilon_i^2 + \gamma_3 \alpha_j \end{cases}$$

$$y_i = I(\quad \beta_0^k + \beta_1 v_i(w_i, k, \epsilon_i^2) + \beta_2 w_i + \theta OP(w_i, \epsilon_i^2) + \epsilon_i^1 + \alpha_j > 0 \quad)$$

and

$$\alpha_j = \lambda_1 \overline{w} + \lambda_2 L_j + u_i$$

where $\overline{w}_j$ is the mean tenant wealth on the landlord's plots of land, L_j are other landlord variables, and u_j is the unobserved component of α_j. One of the problems with this specification is that α_j is only correlated with the observable component of risk aversion (through $\overline{w}$). Unfortunately, it appears we do not have enough data to reliably estimate a more intuitively consistent bivariate matching equation, i.e. $\begin{pmatrix} v_{ij} \\ \alpha_j \end{pmatrix} = f(w_i + \eta_i, \theta)$ (or its inverse, $w_i + \eta_i = f(v_{ij}, \alpha_j, \theta)$).

contract choice equation is normalized to 1. The unobserved component of the random effect does not appear particularly significant. Of particular interest are the coefficients on three landlord variables entering α_j: the occupation dummy variable, the landlord's number of male children, and the landlord's wealth. Presumably being a merchant or notary increases monitoring or measuring costs, while number of children decreases them. Although both variables enter the contract equation with the anticipated sign, neither is significant. The coefficient on landlord wealth in the contract equation is also not significant and fairly small.[45]

7. Conclusions

We have the potential to learn a great deal about markets and economies by studying the determinants of contractual arrangements between parties. For example, empirical evidence on these arrangements can shed light on the functioning of insurance markets, capital markets, or the significance of moral hazard. There has been a great deal of recent empirical work looking at such issues in a wide range of contexts—from historical data, to agriculture in developing countries, to franchising and agriculture in the contemporary U.S..

This paper introduces issues that suggest care in these endeavors. We focus on the empirical implications of potential matching of heterogeneous principals and agents. We argue that there are economic incentives for such matching, and show that when there are econometrically unobserved or partially observed characteristics, this matching is problematic for estimates of optimal contract choice equations.

We then suggest potential solutions for these problems. These solutions revolve around estimation of a matching equation describing how principals and agents match. At the very least, the solutions suggest that a very relevant correlation to examine is the correlation *between* the characteristics the contracting parties. Not only can examination of these "matching equations" solve problems in contract choice equations, but can also provide interesting economic insight on their own. Consideration of these matching equations also points towards future work. If one is willing to make assumptions on the process through which principals and agents match with each other, one can start with basic economic primitives (e.g. utility and production functions) and build a fully speci-

[45] The data apparently do not like landlord wealth entering both the contract choice and matching equations proportionally to the other variables. In this model (i.e. the 4th column without the last coefficient). In this specification, the effect of α_i on contract choice is essentially driven to 0 (the *vines* and *wealth* coefficients are not affected by this). In the full dataset, we do find a significant effect of landlord wealth on contract choice. Interestingly, the sign is positive, i.e. landlord wealth appears to increase the probability of fixed rent contracts. Note that landlord risk-aversion would suggest the reverse, as fixed-rent contracts are less risky for landlords. Perhaps landlord wealth is a better proxy for value of time than the professional variable, or proxies for amount of land holdings which limits ability to monitor or measure.

fied model of both matching and contract choice. Such a model would more fully exploit the interesting information contained in the way principals and agents match.

In our empirical work, we apply our techniques to a historical dataset on agricultural contracts between landlords and tenants in three towns in Renaissance Tuscany. We look for potential evidence of risk-sharing, and of multitasking issues arising from perennial nature of some of the crops. We find a number of interesting empirical results. First, there is very strong evidence that particular types of tenants matched with particular types of landlords. Second, we find that naive estimates ignoring this matching can give misleading results. In our least restrictive specifications, controlling for endogenous matching more than doubles one of the parameters of interest and calls into doubt the very strong significance of the other. Lastly, we end up with some interesting economic conclusions. While most of the literature has not found significant evidence of risk-sharing, we do find evidence for risk-sharing between the peasants and landlords in Tuscany. We also find some evidence that multitasking issues played a role in contractual choice on perennial crops.

These findings are interesting from both a historical point of view and in the context of developing countries. Historians have debated why share contracts were a predominant agrarian arrangements in medieval Tuscany, early modern France and Spain, and the post-bellum U.S. South. At least for medieval Tuscany, we can argue that both risk sharing considerations and multitasking features of medieval viticulture had an impact on the spread of share contracts. In the absence of insurance markets, share contracts seem to have provided an insurance mechanism for tenants while at the same time providing incentives for tenants. Given that some contemporary developing countries have some similar attributes, e.g. imperfect insurance or capital markets, the findings of this paper have the potential of providing interesting insights on current policy debates on the role of agrarian arrangements in these countries.

A. Appendix

We consider identification of the system of reduced form equations (3.5) in the text under the covariance restriction that σ_ϵ^2 is known (e.g. $\sigma_\epsilon^2 = 0$, or that there is no measurement or optimization error in the contract choice equation).:

$$\begin{aligned}
p &= \lambda_0 + \theta w + \gamma\eta + \nu \\
&= \lambda_0 + \theta w + \theta\Phi^{-1}\eta + \nu \\
y &= \lambda_1 + (\beta^p\theta + \pi)w + (\beta^p\gamma + \beta^a)\eta + \beta^p\nu + \epsilon \\
&= \lambda_1 + (\beta^p\theta + \pi)w + (\beta^p\theta + \pi)\Phi^{-1}\eta + \beta^p\nu + \epsilon
\end{aligned}$$

As we consider the case where w, a, and p, are scalars, we have added an explicit constant term to each equation (λ_0 and λ_1). Our objective is to identify the parameters β^p and π. Clearly, Φ^{-1} is not separately identified from σ_η^2, so we combine move Φ^{-1} into σ_η^2 and start with:

$$\begin{aligned}
p &= \lambda_0 + \theta w + \theta\eta + \nu \\
y &= \lambda_1 + (\beta^p\theta + \pi)w + (\beta^p\theta + \pi)\eta + \beta^p\nu + \epsilon
\end{aligned}$$

Coefficients on w in the reduced form equations give us:

$$\widehat{\theta}, (\beta^p\widehat{\theta + \pi})$$

The covariances and variances of the error terms in the 2 equations are estimates of:

$$\begin{aligned}
\widehat{V_{11}} &= \theta^2\sigma_\eta^2 + \sigma_\nu^2 \\
\widehat{V_{22}} &= (\beta^p\theta + \pi)^2\sigma_\eta^2 + \beta^{p^2}\sigma_\nu^2 + \sigma_\epsilon^2 \qquad (A.1) \\
\widehat{V_{12}} &= (\beta^p\theta + \pi)\theta\sigma_\eta^2 + \beta^p\sigma_\nu^2 \qquad (A.2)
\end{aligned}$$

So the question is whether we can use estimates of $\widehat{\theta}, (\beta^p\widehat{\theta + \pi}), \widehat{V_{11}}, \widehat{V_{22}}, \widehat{V_{12}}$ to figure out what β^p and π are. First redefine $\widehat{V_{22}} = \widehat{V_{22}} - \sigma_\epsilon^2$, since σ_ϵ^2 is assumed known (possibly 0). We now have:

$$\begin{aligned}
\widehat{V_{11}} &= \theta^2\sigma_\eta^2 + \sigma_\nu^2 \\
\widehat{V_{22}} &= (\beta^p\theta + \pi)^2\sigma_\eta^2 + \beta^{p^2}\sigma_\nu^2 \qquad (A.3) \\
\widehat{V_{12}} &= (\beta^p\theta + \pi)\theta\sigma_\eta^2 + \beta^p\sigma_\nu^2 \qquad (A.4)
\end{aligned}$$

Solve out each equation for σ_ν^2 and substitute in estimated $(\beta^p\widehat{\theta + \pi})$ and $\widehat{\theta}$ to get:

$$\begin{aligned}
\sigma_\nu^2 &= \widehat{V_{11}} - \widehat{\theta}^2 \sigma_\eta^2 \\
\beta^{p^2} \sigma_\nu^2 &= \widehat{V_{22}} - (\widehat{\beta^p \theta + \pi})^2 \sigma_\eta^2 \qquad \text{(A.5)} \\
\beta^p \sigma_\nu^2 &= \widehat{V_{12}} - (\widehat{\beta^p \theta + \pi}) \widehat{\theta} \sigma_\eta^2 \qquad \text{(A.6)}
\end{aligned}$$

Therefore:

$$\begin{aligned}
\beta^p &= \frac{\widehat{V_{22}} - (\widehat{\beta^p \theta + \pi})^2 \sigma_\eta^2}{\widehat{V_{12}} - (\widehat{\beta^p \theta + \pi}) \widehat{\theta} \sigma_\eta^2} \\
\beta^p &= \frac{\widehat{V_{12}} - (\widehat{\beta^p \theta + \pi}) \widehat{\theta} \sigma_\eta^2}{\widehat{V_{11}} - \widehat{\theta}^2 \sigma_\eta^2}
\end{aligned}$$

and

$$\frac{\widehat{V_{22}} - (\widehat{\beta^p \theta + \pi})^2 \sigma_\eta^2}{\widehat{V_{12}} - (\widehat{\beta^p \theta + \pi}) \widehat{\theta} \sigma_\eta^2} = \frac{\widehat{V_{12}} - (\widehat{\beta^p \theta + \pi}) \widehat{\theta} \sigma_\eta^2}{\widehat{V_{11}} - \widehat{\theta}^2 \sigma_\eta^2}$$

Expanding this gives us:

$$\widehat{V_{22}}\widehat{V_{11}} - (\widehat{V_{22}}\widehat{\theta}^2 + \widehat{V_{11}}(\widehat{\beta^p \theta + \pi})^2)\sigma_\eta^2 + (\widehat{\beta^p \theta + \pi})^2 \widehat{\theta}^2 \sigma_\eta^{2^2} = \widehat{V_{12}}^2 - 2\widehat{V_{12}}(\widehat{\beta^p \theta + \pi})\widehat{\theta}\sigma_\eta^2 + (\widehat{\beta^p \theta + \pi})^2 \widehat{\theta}^2 \sigma_\eta^{2^2}$$

and

$$\widehat{\sigma_\eta^2} = \frac{\left[\widehat{V_{22}}\widehat{V_{11}} - \widehat{V_{12}}^2\right]}{\left[\widehat{V_{22}}\widehat{\theta}^2 - 2\widehat{V_{12}}(\widehat{\beta^p \theta + \pi})\widehat{\theta} + \widehat{V_{11}}(\widehat{\beta^p \theta + \pi})^2\right]}$$

Once we have $\widehat{\sigma_\eta^2}$, obtain:

$$\begin{aligned}
\widehat{\sigma_\nu^2} &= \widehat{V_{11}} - \widehat{\theta}^2 \widehat{\sigma_\eta^2} \\
\widehat{\beta^p} &= \frac{\widehat{V_{12}} - (\widehat{\beta^p \theta + \pi}) \widehat{\theta} \widehat{\sigma_\eta^2}}{\widehat{\sigma_\nu^2}} \\
\widehat{\pi} &= (\widehat{\beta^p \theta + \pi}) - \widehat{\beta^p}\widehat{\theta}
\end{aligned}$$

References

[1] Ackerberg, Daniel, and Botticini, Maristella. “The Choice of Agrarian Contracts in 1427 Tuscany: Risk Sharing, Multitasking, or Capital Market Imperfections?” forthcoming, *Explorations in Economic History*.

[2] Allen, Douglas W., and Lueck, Dean. “Risk Preferences and the Economics of Contracts.” *American Economic Review* 85 (May 1995): 447–51.

[3] ———. “Transaction Costs and the Design of Cropshare Contracts.” *Rand Journal of Economics* 24 (Spring 1993): 78–100.

[4] ———. “Contract Choice in Modern Agriculture: Cropshare versus Cash Rent.” *Journal of Law and Economics* 35 (October 1992): 397–426.

[5] Allen, Franklin. “On the Fixed Nature of Sharecropping.” *Economic Journal* 95 (March 1985): 30–48.

[6] Alston, Lee J. “Tenure Choice in Southern Agriculture, 1930–1960.” *Explorations in Economic History* 18: (April 1981): 211–32.

[7] Alston, Lee J., and Higgs, Robert. “Contractual Mix in Southern Agriculture Since the Civil War: Facts, Hypothesis, and Tests.” *Journal of Economic History* 42 (June 1982): 327-54.

[8] Bandiera, Oriana. “Long-Term Contracts and Share Contracts: Theory and Evidence from 19th-century Sicily.” Working Paper. Boston: Boston College, Department of Economics, 1998.

[9] Bartik, Timothy J. “The Estimation of Demand Parameters in Hedonic Price Models.” *Journal of Political Economy* 95 (February 1987): 81–88.

[10] Becker, Gary. *A Treatise on the Family.* Cambridge, Mass.: Harvard University Press, 1981.

[11] Bell, Clive; Arup, Raha; and Srinivasan, T. N. “Matching and Contractual Performance: Sharecropping in Punjab.” Working Paper. Vanderbilt: Vanderbilt University, April 1995.

[12] Bloch, Marc. *French Rural History.* Berkeley: University of California Press, 1970.

[13] Brandt, Loren, and Hosios, Arthur J. “Credit, Incentives, and Reputation: A Hedonic Analysis of Contractual Wage Profiles.” *Journal of Political Economy* 104 (December 1996): 1172–1226.

[14] Braverman, Avishay, and Stiglitz, Joseph E. "Sharecropping and the Interlinking of Agrarian Markets." *American Economic Review* 72 (September 1982): 695–715 .

[15] Brickley, James A., and Dark, Frederick H. "The Choice of Organizational Form: The Case of Franchising." *Journal of Financial Economics* 18 (July 1987): 401–20.

[16] Carmona, Juan, and Simpson, James. "Vines and Sharecropping: The Case of Cataluña, 1700s–1930s." In *Land, Labor, and Tenure: The Institutional Arrangements of Conflict and Cooperation in Comparative Perspective*, edited by Francesco L. Galassi, Kyle D. Kauffman, and Jonathan Liebowitz. Seville: University of Seville (1998), 87–96.

[17] Cheung, Steven N. *The Theory of Share Tenancy*. Chicago: University Press, 1969.

[18] Chisholm, Darlene C. "Profit-Sharing versus Fixed-Payment Contracts: Evidence from the Motion Pictures Industry." *Journal of Law, Economics, and Organization* 13 (April 1997): 169–201.

[19] Crawford, Vincent P., and Knoer, Elsie Marie. "Job Matching with Heterogeneous Firms and Workers." *Econometrica* 49 (March 1981): 437–50.

[20] Emigh, Rebecca J. "The Spread of Share Cropping in Tuscany: The Political Economy of Transaction Costs." *American Sociological Review* 62 (October 1997): 423–42.

[21] Epple, Dennis. "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products." *Journal of Political Economy* 95 (February 1987): 45–80.

[22] Eswaran, Mukesh, and Kotmal, Ashok. "A Theory of Contractual Structure in Agriculture." *American Economic Review* 75 (June 1985): 352–67.

[23] Ferrall, Christopher, and Shearer, Bruce. "Incentives and Transactions Costs Within the Firm: Estimating an Agency Model Using Payroll Records." IER Working Paper #908. Kingston (Canada): Queen's University, September 1996.

[24] Galassi, Francesco L. "Moral Hazard and Asset Specificity in the Renaissance: The Economics of Sharecropping in 1427 Florence." Manuscript. Warwick: University of Warwick, Department of Economics, 1998.

[25] Galassi, Francesco L.; Mealli, Patrizia; and Pudney, Steve. "Econometrics and the Renaissance: A Discrete Random-Effects Panel Data Model of Farm Tenure in Fifteenth-Century Florence." *Economica* 65 (1999): 535–56.

[26] Galassi, Francesco L., and Kauffman, Kyle D. "The Economics of Draft Animal Choice in Italian Agriculture: A Principal-Agent Approach to the Adoption of an Early Form of Capital." *Journal of European Economic History* 26 (1997): 375–95.

[27] Ghosh, Parikshit. "On the Coexistence of Share, Rent, and Wage Contracts in a Rural Economy." IED Working Paper # 58. Boston: Boston University, 1996.

[28] Grossman, Sanford J., and Hart, Oliver D. "An Analysis of the Principal-Agent Problem." *Econometrica* 51 (January 1983): 7–46.

[29] Hallagan,William. "Self-Selection by Contractual Choice and the Theory of Sharecropping." *Bell Journal of Economics* (Autumn 1987): 344–54.

[30] Hayami,Yujiro, and Otsuka, Keijiro. *The Economics of Contract Choice: An Agrarian Perspective*. Oxford: Oxford University Press, 1993.

[31] Herlihy, David, and Klapisch-Zuber, Christiane. *Les Toscans et leurs familles: un etude du Catasto florentin de 1427*. Paris: Presses de la Fondation Nationale de Sciences Politiques, 1978.

[32] Higgs, Robert. "Patterns of Farm Rental in the Georgia Cotton Belt, 1880–1900." *Journal of Economic History* 34 (June 1974): 468–82.

[33] Hoffman, Philip. "The Economic Theory of Sharecropping in Early Modern France." *Journal of Economic History* 44 (June 1984): 309–19.

[34] Holmstrom, Bengt. "Moral Hazard and Observability." *Bell Journal of Economics* 10 (Spring 1979): 74–91.

[35] Holmstrom, Bengt, and Milgrom, Paul. "Aggregation and Linearity in the Provision of Incentives." *Econometrica* 55 (March 1987): 303–328.

[36] ———. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7 (Spring 1991): 24–45.

[37] ———. "The Firm as an Incentive System." *American Economic Review* 84 (September 1994): 972–91.

[38] Kahn, Shulamit, and Lang, Kevin. "Efficient Estimation of Structural Hedonic Systems." *International Economic Review* 29 (February 1988): 157–166.

[39] Karasker, William S., and Pratt, John W. "Bounding the Effects of Proxy Variables on Regression Coefficients." *Econometrica* 54 (May 1996): 641–55.

[40] Kauffman, Kyle D. "Why Was the Mule Used in Southern Agriculture? Empirical Evidence of Principal-Agent Solutions." *Explorations in Economic History* 30 (1993): 336–51.

[41] Klepper, Steven, and Leamer, Edward E. "Consistent Sets of Estimates for Regressions with Errors in All Variables." *Econometrica* 52 (January 1984): 163–83.

[42] Krueger, Alan. "Ownership, Agency, and Wages: An Examination of Franchising in the Fast Food Industry." *Quarterly Journal of Economics* 106 (February 1991): 75–101.

[43] Laffont, Jean-Jacques, and Matoussi, Mohamed Salah. "Moral Hazard, Financial Constraints, and Sharecropping in El Oulja." *Review of Economic Studies* 62 (April 1995): 381–99.

[44] Lafontaine, Francine. "Agency Theory and Franchising: Some Empirical Results." *RAND Journal of Economics* 23 (Summer 1992): 263–83.

[45] Lafontaine, Francine, and Battacharyya, Sugato. "The Role of Risk in Franchising." *Journal of Corporate Finance* 2 (1995): 39–74.

[46] Lafontaine, Francine, and Slade, Margaret E. "Retail Contracting: Theory and Practice." *Journal of Industrial Economics* 45 (1997): 1–25.

[47] Lanjouw, Jean Olson. "Information and the Operation of Markets: Tests Based on a General Equilibrium Model of Land Leasing in India." Center Discussion Paper NO. 727. New Haven: Yale University, Economic Growth Center, June 1995.

[48] Le- er, Keith B., and Rucker, Randal R. "Transactions Costs and the Efficient Organization of Production: A Study of Timber-Harvesting Contracts." *Journal of Political Economy* 99 (October 1991): 1060–87.

[49] Legros, Patrick, and Newman, Andrew. "Monotone Matching in Perfect and Imperfect Worlds." Working Paper. New York: Columbia University, Department of Economics, 1998.

[50] Luporini, Annalisa, and Parigi, Bruno. "Multi-Task Sharecropping Contracts: The Italian Mezzadria." *Economica* 63 (August 1996): 445–57.

[51] Lyons, Bruce R. "Empirical Relevance of Efficient Contract Theory: Inter-Firm Contracts." *Oxford Reveiw of Economic Policy*, 12 (Winter 1996): 27–52.

[52] Martin, Robert E. "Franchising and Risk Management." *American Economic Review* 78 (December 1988): 954–68.

[53] Milgrom, Paul, and Roberts, John. *Economics, Organization, and Management*. Englewood Cliffs, N. J.: Prentice Hall, 1992.

[54] Mirrlees, James. "Notes on Welfare Economics, Information, and Uncertainty." In *Essays on Economic Behavior Under Uncertainty*, edited by Michael Balch, Daniel McFadden, and S. Wu. Amsterdam: North Holland, 1974.

[55] Mulherin, Harold J. "Specialized Assets, Governmental Regulation, and Organizational Structure in the Natural Gas Industry." *Journal of Institutional and Theoretical Economics* 142 (September 1986): 528–41.

[56] Norton, Seth W. "An Empirical Look at Franchising as an Organizational Form." *Journal of Business* 61 (April 1988): 197–215.

[57] Paarsch, Harry J., and Shearer, Bruce. "Piece Rates, Fixed Wages, and Incentive Effects: Statistical Evidence from Payroll Records." Iowa City: University of Iowa, Department of Economics, February 1997.

[58] ———. "On the Elasticity of Effort for Piece Rates: Evidence from the British Columbia Tree-Planting Industry." Working Paper. Iowa City: University of Iowa, Department of Economics, July 1997.

[59] Reid, John D. "Sharecropping as an Understandable Market Response: The Post-Bellum South." *Journal of Economic History* 33 (April 1973): 106–30.

[60] ———. "Sharecropping and Agricultural Uncertainty." *Economic Development and Cultural Change* 24 (October 1976): 549–76.

[61] ———. "The Theory of Share Tenancy Revisited. Again." *Journal of Political Economy* 85 (April 1977): 403–07.

[62] Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82 (January 1974): 34–55.

[63] Shaban, Radwan Ali. "Testing Between Competing Models of Sharecropping." *Journal of Political Economy* 95 (October 1987): 893–920.

[64] Shapley, Lloyd S., and Shubik, Martin. "The Assignment Game, I: The Core." *International Journal of Game Theory* 1 (1972): 111–30.

[65] Shepard, Andrea. "Contractual Form, Retail Price, and Asset Characteristics." *RAND Journal of Economics* 24 (1993): 58–77.

[66] Shimer, Robert and Smith, Lones. "Assortive Matching and Search." Working Paper. Cambridge, Mass.: MIT, 1996.

[67] Slade, Margaret E. "Multitask Agency and Contract Choice: An Empirical Assessment." *International Economic Review* 37 (1996): 465–86.

[68] State Archives of Florence, Catasto 64, 67, 68, 72, 75, 76, 77, 78, 79, 80, 81, 233, 234, 235, 236, 258, 266, 269.

[69] Stiglitz, Joseph E. "Incentives and Risk Sharing in Sharecropping." *Review of Economic Studies* 41 (1974): 219–55.

Chart 1 - Summary Statistics of Variables

Variable	Description	Mean	S.D.	Min	Max
Rent	=1 if fixed rent contract	0.374	0.484	0	1
Share	=1 if share contract	0.626	0.484	0	1
Vines	=1 if vine crops	0.102	0.302	0	1
Cereals	=1 if cereal crops	0.435	0.496	0	1
Mixed	=1 if both types of crops	0.463	0.499	0	1
vines	$=\begin{cases} 1 \text{ if vines} \\ 0.5 \text{ if mixed} \\ 0 \text{ if cereals} \end{cases}$ (used in regressions)	0.334	0.327	0	1
Tenant wealth	actual tenant wealth in florins	82.51	132.2	0	1552
wealth	normalized ln(tenant wealth) (used in regressions)	0	1	-1.824	2.276
Landlord wealth	normalized	0	1	-1.712	1.769
Occupation	=1 if landlord practices non-agri. occupation	0.683	0.466	0	1
Children	# of adult male children in landlord's household	1.175	1.025	0	5

Source: State Archives of Florence, Catasto 64, 67, 68, 72, 75, 76, 77, 78, 79, 80, 81, 233, 234, 235, 236, 258, 266, 269.

Table 1
All Data

	Share	Fixed Rent	TOTAL
Vines	80.00	12.00	92.00
	(60.72)	(110.66)	(67.23)
Mixed	402.00	16.00	418.00
	(38.29)	(82.37)	(39.98)
Cereals	83.00	309.00	392.00
	(39.79)	(156.08)	(131.45)
TOTAL	565.00	337.00	902.00
	(41.69)	(150.96)	(82.51)

Pescia

	Share	Fixed Rent	TOTAL
Vines	42.00	5.00	47.00
	(88.04)	(134.20)	(92.95)
Mixed	3.00	1.00	4.00
	(23.66)	(344.00)	(103.75)
Cereals	10.00	168.00	178.00
	(110.40)	(184.86)	(180.67)
TOTAL	55.00	174.00	229.00
	(88.60)	(184.32)	(161.33)

Source: See Chart 1
Note: First element in each cell is number of observations, second element (in parentheses) is mean wealth of tenants in that cell.

San Gimignano

	Share	Fixed Rent	TOTAL
Vines	6.00	0.00	6.00
	(78.83)	( )	(78.83)
Mixed	123.00	4.00	127.00
	(56.65)	(42.00)	(56.19)
Cereals	6.00	11.00	17.00
	(32.66)	(165.27)	(118.47)
TOTAL	135.00	15.00	150.00
	(56.57)	(132.40)	(64.16)

Florence

	Share	Fixed Rent	TOTAL
Vines	32.00	7.00	39.00
	(21.46)	(93.85)	(34.46)
Mixed	276.00	11.00	287.00
	(30.27)	(73.27)	(31.91)
Cereals	67.00	130.00	197.00
	(29.89)	(118.10)	(88.10)
TOTAL	375.00	148.00	523.00
	(29.45)	(113.62)	(53.27)

Source: See Chart 1
Note: First element in each cell is number of observations, second element (in parentheses) is mean wealth of tenants in that cell.

Table 2 - Linear Results
Dependent Variable: Contract Choice (=0 if share, =1 if fixed rent).

Parameter	1	2	3	4	5	6	7
	OLS	OLS	Probit	Probit	Fixed Effect	Fixed Effect	OLS
Sample	All	P+S	All	P+S	All	P+S	All
β_0^P	0.8662	0.9252	1.2994	1.5220	-	-	0.9095
(constant)	(0.0226)	(0.0195)	(0.1342)	(0.1577)			(0.0262)
	(0.0000)	(0.0000)	(0.0000)	(0.0000)			(0.0000)
β_0^S	0.4849	0.5011	-0.1847	-0.0770	-	-	0.4735
	(0.0305)	(0.0338)	(0.1845)	(0.2066)			(0.0311)
	(0.0000)	(0.0000)	(0.1583)	(0.3545)			(0.0000)
β_0^F	0.5811	-	0.2877	-	-	-	0.5902
	(0.0230)		(0.0897)				(0.0227)
	(0.0000)		(0.0006)				(0.0000)
β_1	-0.7909	-0.8353	-2.9047	-2.9956	-0.7363	-0.8546	-0.7911
(*vines*)	(0.0453)	(0.0499)	(0.1429)	(0.2622)	(0.0432)	(0.0422)	(0.0450)
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
β_2	0.0897	0.0251	0.4572	0.1322	0.0765	0.0242	-
(*wealth*)	(0.0123)	(0.0142)	(0.0699)	(0.1280)	(0.0128)	(0.0145)	
	(0.0000)	(0.0403)	(0.0000)	(0.1507)	(0.0000)	(0.0447)	
β_2^P	-	-	-	-	-	-	0.0220
							(0.0232)
							(0.1725)
β_2^S	-	-	-	-	-	-	0.0337
							(0.0189)
							(0.0394)
β_2^F	-	-	-	-	-	-	0.1250
							(0.0166)
							(0.0000)
R^2	0.5651	0.7395	-	-	-	-	0.5678

Source: See Chart 1
Notes: Bootstrapped SEs in first parentheses. Bootstrapped p-values in second parentheses. Superscripts indicate town-specific coefficients

Table 3 - Matching Equation and IV Results
Dependent Variables: Columns 1 and 2 - *vines* , Columns 3-7 - Contract Choice

Parameter	1	2	3	4	5	6	7
	Matching Equation	Linear 1st Stage	IV	IV	Fixed Effect IV	Fixed Effect IV	Fix Effect 1st Stage
Sample	All	All	All	P+S	All	P+S	All
β_0^P	-	-	0.8636	0.8924	-	-	-
(constant)			(0.1425)	(0.0893)			
			(0.0008)	(0.0008)			
β_0^S	-	-	0.4803	0.4423	-	-	-
			(0.2482)	(0.1759)			
			(0.0078)	(0.0136)			
β_0^F	-	-	0.5778	-	-	-	-
			(0.1886)				
			(0.0022)				
β_1	-	-	-0.7808	-0.6971	-0.4434	-0.5181	-
(*vines*)			(0.5596)	(0.4053)	(0.4111)	(0.3014)	
			(0.0246)	(0.0362)	(0.0612)	(0.0730)	
β_2	-	-	0.0904	0.0344	0.0954	0.0530	-
(*wealth*)			(0.0367)	(0.0277)	(0.0276)	(0.0229)	
			(0.0272)	(0.0450)	(0.0018)	(0.0048)	
γ_0^P	0.2611	0.3106	-	-	-	-	-
(constant)	(0.0283)	(0.0440)					
	(0.0000)	(0.0000)					
γ_0^S	0.4494	0.4603	-	-	-	-	-
	(0.0168)	(0.0171)					
	(0.0000)	(0.0000)					
γ_0^F	0.3322	0.3323	-	-	-	-	-
	(0.0132)	(0.0133)					
	(0.0000)	(0.0000)					
γ_1	-0.0673	-	-	-	-	-	-
(*wealth*)	(0.0112)						
	(0.0000)						
γ_1^P	-	-0.1381	-	-	-	-	-0.1580
		(0.0398)					(0.0421)
		(0.0012)					(0.0000)
γ_1^S	-	-0.0144	-	-	-	-	-0.0152
		(0.0161)					(0.0168)
		(0.1808)					(0.1832)
γ_1^F	-	-0.0670	-	-	-	-	-0.0534
		(0.0131)					(0.0135)
		(0.0000)					(0.0001)
R^2	0.0971	0.1084	-	-	-	-	-

Source: See Chart 1
Notes: Bootstrapped SEs in first parentheses. Bootstrapped p-values in second parentheses. Superscripts indicate town-specific coefficients

Table 4 - FIML Results - Pescia and San Gimignano Sample

Parameter	1 Probit	2 Ordered Probit First Stage	3 Full Model	4 No functional form Identification
β_o^P	1.5220	-	1.0229	0.1291
(constant)	(0.1481)		(0.2965)	(0.5749)
β_o^S	-0.0770	-	-0.8043	-1.1213
	(0.2313)		(0.4046)	(0.4559)
β_1	-2.9956	-	-1.2274	-2.3640
(*vines*)	(0.2789)		(0.9902)	(0.9564)
β_2	0.1322	-	0.2305	0.3186
(*wealth*)	(0.1356)		(0.1164)	(0.1115)
θ	-	-	-	3.0660
(non-linear wealth)				(1.7632)
γ_1^P	-	-0.4511	-0.4256	-0.4281
(*wealth*)		(0.1127)	(0.1123)	(0.1177)
γ_1^S	-	-0.0929	-0.0956	-0.0984
		(0.1395)	(0.1525)	(0.1585)
$\bar{C}^P$	-	0.7088	0.7149	0.7242
(Upper Cutoff)		(0.1184)	(0.1195)	(0.1164)
$\bar{C}^S$	-	1.8052	1.8073	1.7854
		(0.1494)	(0.1503)	(0.1538)
$\underline{C}^P$	-	0.6451	0.6516	0.6604
(Lower Cutoff)		(0.1159)	(0.1178)	(0.1177)
$\underline{C}^S$	-	-1.1646	-1.1669	-1.1752
		(0.1418)	(0.1399)	(0.1545)
Corr(ε^1,ε^2)	-	-	-0.6587	-0.8611
			(0.3577)	(0.5905)
log likelihood	-90.54	-205.555	-294.31	-293.31

Source: See Chart 1
Notes: Asymptotic SEs in parentheses. Superscripts indicate town-specific coefficients

Table 5 - Random Effects Models - Pescia and San Gimignano Sample

	1	2	3	4
Parameter				
β_o^P	0.2866	0.3614	0.2880	0.3153
(constant)	(0.7387)	(0.8244)	(0.9269)	(0.9453)
β_o^S	-1.0860	-1.0927	-1.1283	-1.0586
	(0.4884)	(0.5040)	(0.6260)	(0.6558)
β_1	-2.2134	-2.1750	-2.3283	-2.3433
(*vines*)	(1.0527)	(1.1235)	(1.1384)	(1.0918)
β_2	0.3128	0.3496	0.3632	0.3564
(*wealth*)	(0.1358)	(0.1588)	(0.1668)	(0.1809)
θ	2.5202	2.3569	2.7553	2.6379
(non-linear wealth)	(2.1030)	(2.3177)	(2.2731)	(2.3730)
γ_1^P	-0.4564	-0.4841	-0.4703	-0.4687
(*wealth*)	(0.1166)	(0.1296)	(0.1303)	(0.1280)
γ_1^S	-0.0965	-0.1302	-0.1272	-0.1217
	(0.1766)	(0.2221)	(0.2179)	(0.1967)
$\bar{C}^P$	0.6952	0.7184	0.7384	0.7611
(Upper Cutoff)	(0.1437)	(0.1470)	(0.2093)	(0.2011)
$\bar{C}^S$	1.8223	1.7977	1.8074	1.7895
	(0.1847)	(0.1993)	(0.2110)	(0.2026)
$\underline{C}^P$	0.6300	0.6530	0.6733	0.6958
(Lower Cutoff)	(0.1433)	(0.1474)	(0.2113)	(0.2038)
$\underline{C}^S$	-1.2191	-1.2613	-1.2272	-1.2577
	(0.1967)	(0.2230)	(0.2199)	(0.2175)
Corr(ε^1,ε^2)	-0.7792	-0.7525	-0.7981	-0.7841
	(0.5824)	(0.5916)	(0.7255)	(0.6545)
γ_3	-0.6884	-0.6449	-0.7049	-0.7531
	(0.6834)	(0.6049)	(0.5540)	(0.6614)
σ_α	0.2621	0.2825	0.2348	0.2049
	(0.2580)	(0.2447)	(0.2154)	(0.2057)
mean tenant		0.0940	0.0786	0.0633
wealth		(0.2018)	(0.1889)	(0.1856)
occupation			0.0753	0.0668
			(0.1903)	(0.1813)
landlord male			-0.0574	-0.0517
children			(0.0965)	(0.0966)
landlord				0.0077
wealth				(0.0931)
landlord wealth				-0.1272
in m.e.				(0.0864)
log likelihood	-292.59	-292.42	-292.02	-289.72

Source: See Chart 1
Notes: Asymptotic SEs in parentheses. Superscripts indicate town-specific coefficients