Essays in Applied Economics

Thèse

Jean-Louis Bago

Doctorat en économique
Philosophiæ doctor (Ph. D.)

Québec, Canada

Essays in Applied Economics

Thèse

Jean-Louis Bago

Sous la direction de:

Bruce Shearer, directeur de recherche
Sylvain Dessy, codirecteur de recherche

Résumé

Cette thèse, structurée en trois essais, aborde des problématiques relatives à l'économie du personnel et l'économie du développement. Mes deux premiers essais étudient de façon expérimentale les explications possibles de l'absence de l'arbitrage entre le risque et les incitations dans les contrats. Mon troisième essai analyse l'impact de la maternité et de l'éducation sur l'emploi autonome des femmes dans les pays en développement.

Dans le premier chapitre, nous analysons la robustesse interne et externe des mesures expérimentales des préférences pour le risque. Les paramètres de préférence pour le risque sont estimés à l'aide d'une série d'expériences de terrain, réalisées dans une entreprise de plantation d'arbres de taille moyenne en Colombie-Britannique. Notre expérience de mesure des préférences pour le risque est construite selon l'approche de Holt et Laury (2002) d'abord avec des loteries à gains faibles puis avec des loteries à gains élevés. Durant l'expérience avec des gains faibles, les participants peuvent gagner entre 2 et 77 dollars. Les montants des loteries sont doublés durant l'expérience avec des gains élevés. Nos résultats montrent que les préférences pour le risque révélées par les travailleurs changent lorsque les montants en jeu augmentent mais la distribution agrégée des préférences pour le risque est restée stable entre les deux traitements. Par la suite, nous analysons la capacité des paramètres de préférence de risque estimés à partir des deux traitements à prédire le choix des travailleurs entre un contrat à la pièce et un contrat à salaire fixe. Les résultats montrent que les préférences pour le risque révélées lorsque les gains sont élevés, prédisent efficacement les choix de contrat. À l'inverse, les préférences pour le risque estimées lorsque les gains sont faibles, ne prédisent pas le choix de contrat. Nous en déduisons donc que l'augmentation des montants en jeu durant l'expérience motive les travailleurs à jouer la loterie plus sérieusement et ainsi révéler leurs préférences réelles pour le risque.

Dans le deuxième chapitre, nous analysons le biais de perception des travailleurs sur la distribution de leur revenu. Une explication possible à l'absence d'arbitrage risque-incitations dans les données contractuelles est que les travailleurs ont une perception incorrecte du risque auquel ils font face dans les contrats. En effet, si les travailleurs sous-estiment le risque, ils seront prêts à travailler dans un environnement très risqué en exigeant une prime de risque plus faible. Ceci réduirait alors le coût de l'incitation pour l'entreprise et supprimerait alors l'arbitrage risque-incitations qui devrait exister dans les contrats. Nous analysons cette problématique en utilisant une série d'expériences de terrain auprès des travailleurs d'une firme de plantation d'arbres en Colombie-Britannique. Au cours de l'expérience, les travailleurs, payés à la pièce, doivent prédire leurs gains espérés sur une durée de 10 jours. À l'aide d'un questionnaire, nous avons obtenu la distribution prédite des gains quotidiens par chaque travailleur. Ensuite, nous construisons la distribution réelle des gains de ces travailleurs en utilisant les données de paie fournies par la firme à l'issue de la saison de plantation. Notre stratégie empirique consiste à apparier la distribution de gains prédite par chaque travailleur avec la distribution réelle des gains pour évaluer l'écart potentiel entre ces deux distributions. Les résultats confirment la présence d'un biais de perception des travailleurs sur leurs gains potentiels sous le contrat de travail à la pièce. En effet, les travailleurs surestiment leurs gains journaliers moyens et sous-estiment la variabilité de leurs gains journaliers et donc le risque auquel ils font face. Par la suite, nos résultats montrent que la sous-estimation du risque accroît la probabilité des travailleurs à choisir un contrat à la pièce au détriment d'un contrat à salaire fixe.

Le troisième chapitre s'intéresse à l'impact de la maternité et de l'éducation sur l'emploi autonome des femmes. En effet, la présence d'enfants en bas âge est souvent désignée dans la littérature comme le déterminant principal de l'emploi autonome des femmes. Toutefois, peu d'évidences empiriques existent dans les pays en développement pour confirmer cette idée. À l'aide de données micro-économiques, nous estimons les effets de la maternité et de l'éducation sur la probabilité d'une femme d'exercer un emploi autonome en Ouganda. Notre stratégie empirique combine l'estimation d'un modèle de type Heckman qui corrige le biais de sélection lié à la participation des femmes sur le marché du travail avec des fonctions de contrôle corrigeant pour l'endogénéité de la fertilité et de l'éducation. Les résultats suggèrent que la présence d'enfants en bas âge n'a pas d'impact significatif sur la probabilité d'une femme d'exercer un emploi autonome en Ouganda. À l'inverse, l'éducation a un effet causal négatif et significatif sur l'emploi autonome. Ce qui implique qu'une meilleure éducation accroît la probabilité pour une femme d'accéder à un emploi salarié sur le marché du travail. En termes de politiques, nos résultats impliquent que l'amélioration de l'éducation des filles est une politique plus efficace pour permettre aux femmes d'accéder à des emplois salariés que les politiques visant la réduction du nombre d'enfants.

Abstract

This thesis consists of three essays, encompassing the fields of development and personnel economics, under the broad banner of applied economics. The first two essays of my thesis contribute to the literature of personnel economics by using multiples field experiments to provide empirical insights to the missing risk-incentives trade-off observed in contractual data. The third essay focuses on the causal effects of motherhood and education on female self-employment in developing countries.

In the first chapter, we investigate the internal and external robustness of risk-preference revealing experiments. We estimate the individual risk preference parameters using field experiments from a medium-sized tree-planting firm in British Columbia. We conducted risk-revealing experiments using the approach of Holt and Laury (2002) successively with low payoff lotteries (low-stakes treatment) and high payoff lotteries (high-stakes treatment). During the low-stakes treatment, workers could win between two 2 and 77 dollars. During the high-stakes treatment, workers could win between 4 and 154 dollars. We find that the aggregate distribution of risk preferences is stable across the two treatments but individual attitudes toward risk change across the low-stake treatment and the high-stake treatment. We explore the ability of the estimated risk preference parameters from the two treatments to predict workers choices between a piece-rate contract and a fixed-wage contract. The results show that the risk preferences measured from the high-stakes treatment predict effectively the contract choice decisions while the risk preference parameters measured from the low-stakes treatment are largely irrelevant. We argue that the increase in stakes led workers to take the lottery more seriously, hence the results from the high-stakes treatment are better measures of their true risk preferences.

In the second chapter, we analyse the presence of risk perception bias among workers who are paid piece rates. A possible explanation for the lack of risk-incentives trade-off in observed contractual data is that workers have biased perceptions of the income risk that they face in contractual settings. For example, if workers underestimate the risk that is present, they will be willing to work in very risky settings for a reduced earnings premium. This would decrease the firm's cost of implementing incentives in risky settings and would suppress the risk-incentives trade-off in observed contractual data. We investigate this issue using daily

payroll data on the earnings of workers who are paid piece rates. We construct the actual distribution of earnings for individual workers in a tree-planting firm. We then elicit each worker's perceived earnings distribution, using a questionnaire. We compare the perceived distribution to the actual distribution. Our results suggest that workers overestimate their average daily earnings and underestimate the standard deviation of their daily earnings and hence the earnings risk that they face. This under-estimation of the risk, increases workers' likelihood of choosing piece rate contracts over the fixed wage contract.

In the third chapter, we present new evidence on the impacts of motherhood and education on women's self-employment probabilities, by accounting for the features of self-employment in a developing country context. Using micro-level data, we estimate the effects of motherhood and education on the self-employment probabilities of women in Uganda. Our estimation framework accounts for selection bias and the endogeneity of motherhood and education jointly, in both the self-employment and the labor force participation regression equations. Consistent estimators of the effects of motherhood and education are obtained by using a Full Information Maximum Likelihood Estimator's method combined with a control function approach for endogenous regressors. We find no evidence of a causal effect of motherhood on women's self-employment probabilities. In contrast, education has a negative causal effect on these probabilities. Both these results differ from the existing literature showing that motherhood, not education, drives women's self-employment probabilities in developed countries. Our results suggest that these findings from the existing literature are not universal, as they do not obtain in the context of a developing economy, where self-employment is predominantly a feature of the insecure informal sector. In this specific context, public policies most effective at reducing the gender gap in pay are likely to be those that pull women out of self-employment, through better education and access to affordable childcare services.

Contents

List of Tables

List of Figures

Dedicated to my mother Elibié Kanmouni and my brother Maurice Bago,

"Greater love has no one than
this, that someone lay down his
life for his friends..."

John 15:13

Acknowledgements

First of all, I give thanks to Almighty God for his help and his guidance throughout my studies. This thesis has been an opportunity for me to learn a lot thanks to the faculty members of the economics department of Laval University and my fellow PhD students.

I would like to express my sincere gratitude to my advisor Prof. Bruce Shearer for his guidance, his kind support, his understanding as well as his caring and parental attitudes towards me throughout the course of my thesis. Thank you for your patience, your encouragement and for always pushing me to strive harder. I was extremely lucky to have you as my advisor. I wish to extend my deep gratitude to my co-supervisor Prof. Sylvain Dessy for accepting to collaborate on the third paper of my thesis. I thank you for providing me with invaluable advice and constant inspiration. I am also indepted to Prof. Bernard Fortin for his insightful comments on my research on various occasions. Many thanks to my thesis committee members Luca Tiberti and Sonia Laszlo for reading my thesis and generously providing me with very helpful comments and recommendations.

During my thesis, I have spent one year of internship at the Ministry of Finance of Quebec. I would like to thank my internship supervisor Francis Hebert and my co-worker Imad Rherrad.

I am also grateful to my fellow PhD students Blanchard Conombo, Koffi Akakpo, and Ibrahima Sarr for discussions and invaluable inputs to my dissertation. I also thank my friends and colleagues from Burkina Faso and Quebec who have always been there for me during my journey in Quebec, especially Francois d'Assises Bationo, Millogo Abdoulaye, Daouda Belem, Jean-Baptiste Ouédraogo, Boukary Traoré, Benoit Kafando, Eli Sawadogo, Ousmane Z. Traoré, Yacin Ouedraogo, Aude Koutaba, Aristide Valéa, Alou Dembélé, Moussa Ouédraogo, Pouiré Ouédraogo, Gautier Yé, Ernest Ouédraogo, Minata W. Souratié, Aimé David Zoundi, Finagnon Antoine Dedewanou, Morvan Nongni Donfack, Elysee Aristide Houndetoungan, Rolande Kpekou Tossou, Morvan Nongui and Guy Morel Kossivi Amouzou Agbe. I am also indepted to Mrs Beatriz Sifuentes and her family for their warm welcome at my first arrival in Canada. Special thanks to Miaba Louise Lompo for her support.

Last but not least, I dedicace this thesis to my mother Elibié Kammouni and my brother Maurice Bago for their unconditional love, sacrifice, financial support and prayers.

Foreword

The three chapters of this thesis are separate articles submitted or in preparation for submission to peer-reviewed scientific journals. The first two chapters were written jointly with my thesis supervisor Prof. Bruce Shearer, full professor in the Department of Economics at Laval University. The third paper is written with Prof. Sylvain Dessy, also full professor in the Department of Economics at Laval University. I am the principal author of the three chapters presented in this thesis. The first essay entitled "Les Mesures Expérimentales des Préférences pour le Risque prédisent- elles le choix des contrats? Evidence d'une expérience de terrain" and the second essay titled "Perception Bias and Risk-Incentives Trade-off : Evidence from Field Experiments" are in preparation to be submitted to peer-reviewed scientific journals. The third essay entitled "Motherhood and Female Self-Employment: Theory and Evidence from Uganda " has been submitted to Journal of Labor Economics.

Introduction

Decision-making process in the workplace is a subject of great interest to economists. In the labor market, individuals make decisions in uncertain environments based on their observed and unobserved characteristics such as their risk preferences, their risk perception, their ability, their education and the demographic characteristics of their households. My PhD thesis explores theoretically and empirically the role of these individual characteristics on worker's contract choices and employment outcomes in three essays in applied economics.

Risk-incentives trade-off plays a central role in economic models of incentive contracts. Providing incentives typically implies forcing workers to share risk, which imposes costs on risk-averse workers, and raises contracting costs (Prendergast, 1999). Piece-rate contracts attach workers' pay to their productivity, exposing them to income risk. Risk-averse workers must be compensated for working in risky environments which increases the cost of assigning such workers to piece rate contracts. The principal-agent model therefore predicts that piece rate contracts will be avoided in risky settings (Hölmstrom, 1979). Yet, the empirical role of risk in determining contracts is less clear-cut (Allen and Lueck, 1992; Prendergast, 2000). Empirical models suffer from the fact that risk and risk preferences are unobserved, complicating the testing of theoretical comparative static results. The first two chapters of my thesis contribute to the literature of personnel economics by using field experiments to provide empirical insights to the missing risk-incentives trade-off observed in contractual data. Both essays used the same experimental data to investigate two alternative explanations of the missing risk-incentives tradeoff.

The first possible explanation to the missing trade-off concerns difficulties in observing risk preferences. In fact, the risk-incentives trade-off depend on a worker's risk preferences: risk averse workers require a risk premium to accept risky contracts while risk neutral or risk loving workers require no premium. Since risk preferences are not observed, many researchers use proxies for risk preferences to identify the risk-incentive trade-off, leading to mixed results (Allen and Lueck, 1999; Ackerberg and Botticini, 2002). An alternative approach is to use experimental measures of risk preferences (Holt and Laury, 2002; Bellemare and Shearer, 2010) and incorporate them into contractual data. Bellemare and Shearer (2013) take such an approach, combining risk-preference revealing experiments with payroll data to identify and estimate a structural principal-agent model. The validity of such an approach depends, inter alia, on the ability of the experiments to identify risk preferences. We address this issue in the first chapter by considering the ability of the experimentally identified risk preferences to predict contractual choices in real-world settings. To do so, we exploit multiple field experiments, conducted within a real firm. Our experiments took place within a tree-planting firm, located in British Columbia, Canada. Workers in this firm are hired to plant trees on recently logged parcels of land and are typically paid piece rates – daily earnings are the product of the piece rate and the number of trees planted. Since the number of trees a worker can plant depends on elements that are beyond his/her control (such as weather, the slope and the hardness of the ground) workers are exposed to daily income risk under a piece-rate contract. Fixed-wage contracts eliminate that risk to workers. In the first set of experiments, we conducted a risk-preference-revealing experiment and measured the ability of the estimated risk preferences parameters to predict workers contract choices. We repeat this risk revealing experiments twice on each worker: once at low-risk stakes and once at high-risk stakes. We then use the measured risk preferences to predict worker behaviour in a second set of experiments, in which workers choose between piece-rate and fixed-wage contracts. The results indicate that the risk preferences measured from the high-stakes treatment predict effectively the contract choice decisions while the risk preference parameters measured from the low-stakes treatment are largely irrelevant. We argue that the increase in stakes led workers to take the lottery more seriously, hence the results from the high-stakes treatment are better measures of their true risk preferences.

The second explanation to the lack of risk-incentives trade off is the perception bias. In fact, if workers underestimate the risk that is present, they will be willing to work in risky settings for a much reduced earnings premium. This would reduce the firm's cost of implementing incentives in risky settings and would suppress the risk-incentives trade-off in observed contractual data. In the second chapter, we address this issue by considering workers' perception on their earnings risk within a real firm. To do so, we exploit data from field experiments conducted in a tree planting firm. First, we elicit each worker's perceived earnings distribution, using two questions. The first question is a single value forecasting, in which workers were asked to predict their average earnings. The second question is a probabilistic forecasting. Workers were asked to predict the probability that their earnings would lie within a number of fixed intervals. We use the worker's responses to construct their perception of their earnings distribution. Subsequently, we construct the actual distribution of earnings for each workers based on their payroll records and observed daily earnings. We compare the perceived distribution to the actual distribution in order to evaluate the presence of perception bias on earnings risk and its effects on workers risk-incentive trade-off in incentive contracts. Overall, we find evidence that workers underestimate the risk they face under the piece rate contract and overestimate their average earnings. Subsequently, we analyzed the ability of the perception bias to predict workers contract choice. We find that the greater is the degree of underestimation, the more likely workers are to select piece rates.

The third empirical issue, that we addressed in this thesis, is related to the importance of fertility and education as determinants of women's self-employment in developing countries. The comparison of male and female earnings is a subject of great interest to economists. It is well documented that women earn less than men (Ginther and Hayes, 1999; Bayard et al., 2003; Blackaby et al., 2005; Aizer, 2010). Yet, the causes of this pay gap are a subject of on-going research. Understanding why women still earn a lot less than men is crucial in guiding policy actions aimed at reducing the gender pay gap. In developing countries, much of the problem seems to be due to the concentration of women in low-paying sectors of the economy, often characterized by informal self-employment. Recent researches have emphasized the importance of motherhood (and particularly the presence of young children) as a determinant of women being self-employed (Connelly, 1992; Bianchi, 2000; Wellington, 2006; Boserup et al., 2013; Lim, 2018; Semykina, 2018). Self-employment is presented as an occupational strategy allowing women to balance work and family commitments (Connelly, 1992; Craig et al., 2012). Yet, direct empirical evidence supporting this proposition is largely limited to developed countries, where women's self-employment probabilities are lower than those of men, and are much smaller than for developing countries. Moreover, these findings are also being obtained with empirical strategies that overlook the interrelation of education and self-employment decisions. In this paper, we present new quantitative evidence on the importance of motherhood to women's self-employment in a developing economy's context, using micro-level data from Uganda. To guide our empirical analysis, we develop a theoretical model explaining how a woman's individual characteristics affect her labor force participation and self-employment decisions. Our model highlights a key distinguishing feature of self-employment in a developing economy's context: balancing paid work and unpaid child care labor in self-employment may enhance child quality compared to wage employment, when caregiving to younger children is primarily a mother's responsibility. However, it also entails a motherhood penalty in the form of loss of labor productivity. This loss of productivity has two different sources. First, there is the fact that self-employment takes place in the informal economy characterized by insecure property rights, which discourages the adoption of productivity-enhancing technologies, thus leading to loss of labor productivity relative to wage employment—a formal sector activity. Second, in self-employment women work while caring for their younger children. Yet caring for younger children while working also reduces a mother's labor productivity. Since wage employment, which takes place predominantly in the formal sector, also entails a motherhood penalty in the form of childcare expenses, having younger children therefore has an ambiguous effect on women's self-employment decisions. This implies that whether children matter to women's self-employment probabilities or not is an empirical issue. Our empirical strategy accounts for nonrandom self-selection and the endogeneity of multiple regressors jointly. We find no evidence of a causal effect of motherhood on a woman's self-employment. Instead, we find that an additional year of schooling completed reduces a woman's probability of self-employment by 3.05 percent. In terms of policy implication, these findings suggest that the sociocultural and institutional contexts determine the nature of factors that drive women's self-employment probabilities. In a developing economy, participation in self-employment is more likely to be involuntary than voluntary for women, because it takes place in the informal sector characterized by insecure property rights (Heintz and Valodia, 2008). To reduce the gender gap in pay in this context, public policy is likely to be more effective if it focuses on promoting female education as well as households' access to quality professional childcare services. Education would draw women out of the informal sector and into the formal sector where the bulk of wage employment takes place, while affordable childcare services would reduce the cost of formal sector employment for mothers.

Bibliography

Ackerberg, D. A. and M. Botticini (2002): "Endogenous matching and the empirical determinants of contract form," *Journal of Political Economy*, 110, 564–591.

Aizer, A. (2010): "The gender wage gap and domestic violence," *American Economic Review*, 100, 1847–59.

Allen, D. and D. Lueck (1992): "Contract choice in modern agriculture: cash rent versus cropshare," *The Journal of Law and Economics*, 35, 397–426.

Allen, D. W. and D. Lueck (1999): "The role of risk in contract choice," *Journal of Law, Economics, and Organization*, 15, 704–736.

Bayard, K., J. Hellerstein, D. Neumark, and K. Troske (2003): "New evidence on sex segregation and sex differences in wages from matched employee-employer data," *Journal of labor Economics*, 21, 887–922.

Bellemare, C. and B. Shearer (2010): "Sorting, incentives and risk preferences: Evidence from a field experiment," *Economics Letters*, 108, 345–348.

——— (2013): "Multidimensional heterogeneity and the economic importance of risk and matching: evidence from contractual data and field experiments," *The RAND Journal of Economics*, 44, 361–389.

Bianchi, S. M. (2000): "Maternal employment and time with children: Dramatic change or surprising continuity?" *Demography*, 37, 401–414.

Blackaby, D., A. L. Booth, and J. Frank (2005): "Outside offers and the gender pay gap: Empirical evidence from the UK academic labour market," *The Economic Journal*, 115, F81–F107.

Boserup, E., S. F. Tan, and C. Toulmin (2013): *Woman's role in economic development*, Routledge.

Connelly, R. (1992): "Self-employment and providing child care," *Demography*, 29, 17–29.

Craig, L., A. Powell, and N. Cortis (2012): "Self-employment, work-family time and the gender division of labour," *Work, employment and society*, 26, 716–734.

Ginther, D. K. and K. J. Hayes (1999): "Gender differences in salary and promotion in the humanities," *American Economic Review*, 89, 397–402.

Heintz, J. and I. Valodia (2008): "Informality in Africa: A review," *Background paper for the Swedish International Development Cooperation Agency (Sida), Unpublished Working Paper.*

Hölmstrom, B. (1979): "Moral hazard and observability," *The Bell journal of economics*, 74–91.

Holt, C. A. and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *The American Economic Review*, 92, 1644–1655.

Lim, K. (2018): "Do American mothers use self-employment as a flexible work alternative?" *Review of Economics of the Household*, 1–38.

Prendergast, C. (1999): "The provision of incentives in firms," *Journal of economic literature*, 37, 7–63.

——— (2000): "What trade-off of risk and incentives?" *The American Economic Review*, 90, 421–425.

Semykina, A. (2018): "Self-employment among women: Do children matter more than we previously thought?" *Journal of Applied Econometrics*, 33, 416–434.

Wellington, A. J. (2006): "Self-employment: the new solution for balancing family and career?" *Labour Economics*, 13, 357–386.

Chapter 1

Les mesures expérimentales des préférences pour le risque prédisent-elles le choix des contrats? Evidence d'une expérience de terrain.

1.1 Résumé

Dans cet article, nous analysons la robustesse interne et externe des mesures expérimentales des préférences pour le risque. Les paramètres de préférence pour le risque sont estimés à l'aide d'une expérience de terrain réalisée dans une firme de plantation d'arbres en Colombie-Britannique. L'expérience utilisée pour mesurer les préférences pour le risque des travailleurs est construite selon l'approche de Holt et Laury (2002). Nous effectuons l'expérience d'abord avec des loteries à gains faibles puis avec Des loteries à gains élevés. Durant l'expérience avec des gains faibles, les participants peuvent gagner entre 2 et 77 dollars. Les montants sont doublés durant l'expérience avec des gains élevés. Nos résultats montrent que les préférences pour le risque révélées par les travailleurs changent lorsque les montants en jeu augmentent mais la distribution agrégée des préférences pour le risque est restée stable entre les deux traitements. Par la suite, nous analysons la capacité des paramètres de préférence pour le risque à prédire le choix des travailleurs entre un contrat à la pièce et un contrat à salaire fixe. Les résultats montrent que les préférences pour le risque révélées lorsque les gains possibles des loteries sont élevés, prédisent efficacement les choix de contrat. À l'inverse, les préférences pour le risque estimées lorsque les gains sont faibles ne prédisent pas le choix de contrat. Nous en déduisons donc que l'augmentation des montants en jeu durant l'expérience motive les travailleurs à jouer la loterie plus sérieusement et ainsi révéler leurs préférences réelles pour le risque.

Mots-clés: préférence pour le risque, expérience de terrain, contrat à la pièce, salaire fixe.
Codes JEL: C93, D86, J33.

1.2 Abstract

This paper investigates the internal and external robustness of risk-preference revealing experiments. We estimate the individual risk preference parameters using field experiments from a medium-sized tree-planting firm in British Columbia. We conducted risk-revealing experiments using the approach of Holt and Laury (2002) successively with low payoff lotteries (low-stakes treatment) and high payoff lotteries (high-stakes treatment). During the low-stakes treatment, workers could win between two 2 and 77 dollars. During the high-stakes treatment, workers could win between 4 and 154 dollars. We find that the aggregate distribution of risk preferences is stable across the two treatments but individual attitudes toward risk change across the low-stake treatment and the high-stake treatment. We explore the ability of the estimated risk preference parameters from the two treatments to predict workers choices between a piece-rate contract and a fixed-wage contract. The results show that the risk preferences measured from the high-stakes treatment predict effectively the contract choice decisions while the risk preference parameters measured from the low-stakes treatment are largely irrelevant. We argue that the increase in stakes led workers to take the lottery more seriously, hence the results from the high-stakes treatment are better measures of their true risk preferences.

Keywords: risk preference, field experiment, piece rate, fixed-wage.
JEL Codes: C93, D86, J33.

1.3 Introduction

Le risque et les préférences pour le risque occupent une place centrale dans les modèles économiques des contrats incitatifs. Au plan théorique, en imposant un partage de risque entre la firme et ses travailleurs, les incitations génèrent des coûts de risque pour les travailleurs qui sont averses au risque. Ceci suscite des coûts d'implémentation des contrats incitatifs pour la firme dans un environnement risqué (Prendergast, 1999). Cependant, le rôle du risque dans la détermination des contrats reste peu clarifié dans la littérature empirique (Prendergast, 2000). Le défi majeur des études empiriques est l'inobservabilité des préférences pour le risque qui complique toutes les analyses visant à tester les prévisions théoriques des statistiques comparatives sur les contrats. Pour surmonter cette difficulté, plusieurs auteurs font recours à des proxies pour mesurer les préférences pour le risque et utilisent ces proxies pour expliquer la forme des contrats, mais les résultats restent mitigés (Allen and Lueck, 1999; Ackerberg and Botticini, 2002). Une approche alternative est l'utilisation de mesures expérimentales des préférences pour le risque (Holt and Laury, 2002; Cadsby et al., 2007; Eckel and Grossman, 2008; Bellemare and Shearer, 2010; Dohmen and Falk, 2011; Dohmen et al., 2011) conjointement avec les données relatives à la forme des contrats. Bellemare and Shearer (2013) utilisent une telle approche pour identifier les préférences pour le risque et estimer un modèle structurel d'agence. Ils utilisent leur modèle pour évaluer les gains potentiels pour la firme d'apparier les travailleurs au niveau de risque selon leurs préférences.

La validité d'une telle approche dépend du potentiel des expériences à identifier correctement les préférences pour le risque et de la pertinence de cette mesure expérimentale dans un contexte réel. Cette forme de validité externe ayant des implications importantes justifiera l'incorporation des résultats de l'expérience dans l'environnement des contrats.

Nous abordons cette question en considérant la capacité des préférences pour le risque mesurées de façon expérimentale à prédire des choix des contrats des travailleurs dans un environnement réel. Nous utilisons une séquence d'expériences de terrain, menées dans une firme réelle. Dans la première série d'expériences, nous mesurons les préférences de risque des travailleurs en utilisant les méthodes de Holt and Laury (2002). Ces expériences sont répétées deux fois sur chaque travailleur. La première expérience utilise des loteries avec des mises faibles et la seconde utilise des loteries avec des mises élevées. Les préférences pour le risque mesurées sont ensuite utilisées pour prédire le comportement des travailleurs dans une deuxième expérience, dans laquelle les travailleurs choisissent entre les contrats à la pièce et les contrats à salaire fixe. Il est important de noter que ces choix de contrats conduisent à des conséquences réelles: les participants sont payés selon le contrat résultant de leurs décisions pendant une période de deux jours de travail.

Nos expériences ont été réalisées entre Mai et Juin 2016 dans une entreprise de plantation d'arbres, située en Colombie-Britannique, au Canada. Les travailleurs de cette entreprise sont embauchés pour planter des arbres sur des parcelles récemment déboisées et sont généralement payés selon des contrats à la pièce. Les gains journaliers de chaque travailleur sont les taux à la pièce multipliés par le nombre d'arbres plantés. Étant donné que le nombre d'arbres que le travailleur peut planter dépend de son effort, mais aussi de facteurs hors de son contrôle (comme la météo, la pente et la dureté du sol...), les travailleurs sont exposés au risque de revenu quotidien sous le contrat à la pièce. Les contrats à salaire fixe éliminent ce risque pour les travailleurs.

Pour mesurer les préférences pour le risque, nous utilisons les méthodes expérimentales développées par Holt and Laury (2002). L'expérience consiste pour les travailleurs à prendre 10 décisions consécutives dont chacune consiste à choisir entre une loterie à risque faible et une loterie à risque élevé[1]. La loterie moins risquée paye soit 40$, soit 32$ alors que la loterie plus risquée paye soit 77$, soit 2$. Nous dénotons cette expérience le traitement avec des loteries à mises faibles (LMF). Les 10 décisions se distinguent par la probabilité de gains de montant élevé (40$ pour la loterie moins risquée et 77$ pour la loterie plus risquée). À la première décision, cette probabilité est fixée à 0,1 pour chaque loterie. Étant donné la probabilité faible de gagner, seul un agent ayant une grande préférence pour le risque choisira la loterie plus risquée à cette décision. La probabilité de gagner s'accroît de 0,1 à chaque décision, atteignant ainsi la probabilité certaine de 1 à la dixième décision. Lorsqu'un participant a effectué toutes ses décisions, l'une d'entre elles a été tirée au sort et la loterie choisie par le participant à cette décision a été jouée pour déterminer son gain réel qui lui fut payé. Ensuite, les mises ont été doublées dans le deuxième traitement de sorte que la loterie moins risquée paye entre 80$ et 64$ et la loterie plus risquée paye entre 154$ et 4$. Nous dénotons cette expérience le traitement avec des loteries à mises élevées (LME).

Durant la seconde expérience, les travailleurs doivent choisir à douze reprises entre leur contrat à la pièce régulier et une série de contrats à salaire fixe. Les contrats à salaire fixe commencent à 100$ pour la journée et augmentent par 50$ à chaque décision pour atteindre 700$ par jour. À chaque décision le travailleur indique son contrat préféré. Lorsque les douze décisions sont effectuées, l'une d'entre elles a été tirée aléatoirement et le contrat choisi par le travailleur à cette décision lui sera implémenté pour les deux prochains jours de travail.

Nous développons un modèle de choix optimal de contrat en fonction de la préférence pour le risque et de l'habileté du travailleur. Nous dérivons à partir de ce modèle l'équivalent certain du contrat à la pièce pour le travailleur. Ceci représente le contrat à salaire fixe qui procure la même utilité au travailleur que son contrat à la pièce. Nos statiques comparatives montrent que l'équivalent certain est une fonction décroissante du degré d'aversion au risque du travailleur et une fonction croissante de son habileté. Nous dérivons donc des prédictions théoriques qui peuvent être testées en utilisant les données expérimentales et les données de plantation de la firme. Tout d'abord, conditionnellement à l'habileté, les travailleurs plus averses au risque devraient être plus enclins à accepter le contrat à salaire fixe. Deuxièmement, étant donné les préférences de risque, les travailleurs hautement habiles devraient préférer davantage les contrats à la pièce.

En utilisant les données provenant de nos expériences, nous testons la validité de ces prédictions théoriques. L'analyse de l'effet de l'échelle de la loterie sur la mesure des préférences pour le risque révèle que la distribution agrégée des préférences pour le risque est stable de l'expérience LMF à l'expérience LME. Cependant, les préférences pour le risque mesurées pour un individu donné changent lorsque la taille des mises est doublée. En testant les prédictions théoriques, les résultats montrent que les mesures de préférence pour le risque provenant de l'expérience LME prédisent bien le choix de contrats des travailleurs. Les travailleurs les plus averses au risque ont démontré une plus faible prévalence à choisir le contrat à la pièce. Les travailleurs les plus habiles s'auto-sélectionnent davantage dans les contrats à la pièce que dans les contrats à salaire fixe. À l'inverse, les mesures obtenues par l'expérience LMF ne prédisent pas le choix des contrats. Nous soutenons alors que l'augmentation de

[1]Le choix de la loterie moins risquée sera qualifié de "choix-sûr" en référence à la loterie plus risquée.

la taille des mises incite les travailleurs à prendre les loteries plus au sérieux et à révéler ainsi leurs préférences plus efficacement.

Nos travaux rejoignent ceux de Belzil and Sidibé (2016) qui ont analysé la validité interne et externe de ces mesures expérimentales des préférences pour le risque et le facteur d'escompte du temps dans le financement de l'éducation. En utilisant une expérience de choix entre un paiement cash immédiat et une bourse échelonnée, ils montrent que les facteurs d'escompte de long terme et la perception du risque concernant le paiement prédisent bien leur décision de participation ou non à la bourse échelonnée des participants. À l'opposé, les estimateurs de court terme n'expliquent pas la prise de décision.

La suite du papier est organisée comme suit: la section 2 présente les détails institutionnels relatifs à l'industrie forestière et à la politique de reforestation de la Colombie-Britannique. La troisième section s'intéresse au protocole expérimental utilisé. Une description des données utilisées est faite dans la section 4. La section 5 est consacrée au modèle structurel et aux prédictions théoriques concernant le choix des contrats. La section 6 analyse la validité interne des mesures expérimentales des préférences pour le risque obtenues avec l'approche de Holt and Laury (2002). La section 7 teste la capacité des préférences pour le risque mesurées à prédire le choix des contrats conformément aux prédictions théoriques. Enfin, la dernière section effectue une discussion des résultats et conclut l'étude.

1.4 Les détails institutionnels

Cette section présente l'industrie et la politique de reforestation de la Colombie- Britannique. Elle donne également un aperçu des firmes et de l'activité de plantation d'arbres.

1.4.1 L'industrie forestière et la politique de reforestation en Colombie-Britannique

La province de Colombie-Britannique, située sur la Côte-Ouest du Canada est l'une des plus grandes exportatrices mondiales de bois d'oeuvre. L'exploitation et le commerce du bois d'oeuvre en Colombie-Britannique représentent le quart de la production nord-américaine et plus de la moitié de la production canadienne. Le secteur forestier est un levier important de l'économie de la Colombie britannique. En 2015, ce secteur a généré 65 500 emplois soit 7,7% des emplois dans la province et a contribué à hauteur de 3,4% au Produit Intérieur Brut[2].

Deux types de firmes opèrent dans le secteur forestier en Colombie-Britannique. En amont se trouvent les firmes spécialisées d'exploitation forestière qui achètent des licences du gouvernement pour la coupe du bois à des fins industrielles et en aval, les firmes de reforestation chargées du reboisement des surfaces exploitées. Pour assurer sa production régulière de bois, la Colombie-Britannique dispose d'une politique de reforestation active. Les espaces de forêts coupées durant l'année pour alimenter l'industrie du bois sont toujours reboisés dans les saisons suivant la fin de la coupe. Contrairement à la régénération naturelle de la forêt qui est très lente et non contrôlable, cette reforestation active permet d'accélérer le processus de restauration des zones forestières détruites par la coupe, mais aussi

[2]Référence: http://www.bcfii.ca/bc-forests-and-markets/bc-forest-sector-overview

de protéger la biodiversité. Les deux acteurs principaux impliqués dans la politique de reforestation sont le Ministère des forêts, des terres et des ressources naturelles (MFTR) et les firmes de reboisement.

Les contrats de reboisement sont attribués aux entreprises selon un processus compétitif d'enchères publiques sous pli fermé. Durant la période d'automne précédant chaque saison de plantation allant du printemps à l'été, le MFTR lance un appel à soumission aux firmes de reboisement pour planter des arbres dans les zones récemment exploitées. Avant de faire son offre, chaque firme effectue une étude du terrain à planter pour estimer les coûts qu'elle devra supporter pour compléter le contrat de plantation étant donné les conditions du terrain. À la suite de son étude du terrain, la firme soumet son offre en précisant son prix par arbre. La firme ayant proposé le prix de soumission le plus faible remporte le contrat. Dès lors qu'un contrat est signé entre la firme et le MFTR, la firme doit embaucher des planteurs d'arbres pour reboiser le terrain qui lui est attribué. Typiquement, les travailleurs sont payés selon un salaire journalier à la pièce. Dans ce contrat, le gain du travailleur est entièrement proportionnel à sa production mesurée en nombre d'arbres plantés. Cependant, des contrats à salaire fixe sont parfois convenus entre la firme et l'employé. Les firmes restent tenues de respecter le salaire minimum exigé en vertu de la loi[3].

1.4.2 La firme et l'activité de plantation d'arbres

Les données utilisées dans ce papier proviennent de la saison de plantation de 2016 et ont été récoltées dans une firme de taille moyenne de plantation d'arbres localisée en Colombie-Britannique. L'activité de plantation d'arbres est saisonnière et se déroule chaque année du printemps à l'été.

La firme recrute pendant chaque saison un groupe de travailleurs. On y retrouve les travailleurs permanents et les travailleurs temporaires[4]. Les terrains à planter sont subdivisés en parcelles dont la taille peut varier de 10 à 15 travailleurs. À chaque dizaine de planteurs est attribué un superviseur dont la tâche est de contrôler la qualité du travail et d'approvisionner les travailleurs en plants.
Chaque matin, les travailleurs reçoivent une pelle spécialement conçue pour la plantation et un sac d'arbres à planter. La tâche de planteur d'arbres est relativement simple, mais minutieuse et nécessite une bonne condition physique. Pour planter un arbre, on creuse d'abord un trou dans le sol assez large et profond pour pouvoir étaler toutes les racines de l'arbre. Ensuite, on positionne verticalement le plant dans le trou. Enfin, il faut remplir le trou par couches successives avec la terre qui a été mise de côté en veillant à bien recouvrir toutes les racines de l'arbre de terre. Des règles relatives à la distance entre les plants, la profondeur et la largeur des trous doivent être respectées pour garantir une pousse adéquate des arbres. Si ces règles ne sont pas respectées, la firme s'expose à des amendes du gouvernement qui effectue également des contrôles. La probabilité de survie d'un arbre dépend de la qualité de sa plantation.

Quand les travailleurs sont payés à la pièce, le salaire journalier de chaque travailleur dépend exclusivement de sa production mesurée par le nombre d'arbres plantés et du taux à la pièce. Il n'y a aucun salaire de base associé au contrat à la pièce. Or, le nombre d'arbres plantés sur une parcelle par un travailleur donné est fonction de son effort, mais aussi des conditions naturelles aléatoires du terrain.

[3]Nous avons exclus de l'échantillon deux observations correspondantes à des gains inférieurs au salaire minimum garanti en Colombie-Britannique afin de ne pas prendre en compte les planteurs ayant travaillé une demi-journée par exemple

[4]Les travailleurs temporaires sont en majorité des étudiants recherchant un travail d'été.

Un sol sableux et plat est plus facile à planter qu'un terrain rocailleux, abrupt et non préparé. La Colombie-Britannique étant une région montagneuse, les terrains présentent de grandes variabilités. En conséquence, pour un même taux à la pièce, le planteur préfère travailler sur une parcelle facile à planter, car pour un niveau d'effort donné, il peut planter plus d'arbres et gagner plus d'argent. Ainsi pour prendre en compte les conditions du terrain, les taux à la pièce diffèrent donc d'une parcelle à l'autre. Pour chaque parcelle, la firme fixe le taux à la pièce en fonction du nombre d'arbres qu'un travailleur peut espérer planter par jour et de la masse salariale prévue par la firme. Pour inciter les travailleurs à planter sur un terrain difficile, le taux à la pièce sera plus élevé afin de satisfaire sa contrainte de participation. Le gain du travailleur sous un contrat à la pièce est donc soumis au risque lié aux conditions du terrain.

Étant donné qu'il n'y a pas de travail en équipe [5], le gain de chaque planteur est indépendant de celui des autres planteurs. Ce qui permet d'observer l'effet individuel des incitatifs monétaires sur le comportement des travailleurs (Shearer, 2004).

1.5 Le protocole expérimental

Dans le cadre de ce papier, deux expériences de terrain ont été réalisées dans la firme durant la période de mai à juin 2016 afin de mesurer la préférence pour le risque et sa capacité à prédire le choix de contrat du travailleur entre son contrat à la pièce et un contrat à salaire fixe.

1.5.1 L'expérience pour mesurer les préférences pour le risque

La première expérience, effectuée avec 63 travailleurs, vise à mesurer les préférences pour le risque des travailleurs. Cette expérience qui s'inspire du design expérimental de Holt and Laury (2002) s'est déroulée sur trois jours du 25 au 27 mai 2016. Chaque matin, l'équipe des chercheurs a rencontré les travailleurs avant leur départ sur le site de reboisement. En se présentant comme des économistes pratiquant une étude de terrain sur le comportement des travailleurs en présence de risque, les chercheurs ont invité les planteurs à participer à cette expérience de façon volontaire. Chaque participant est assuré de recevoir une compensation initiale forfaitaire de 20$ pour sa participation en plus des gains supplémentaires qui varient entre 2$ et 77$ en fonction de ses actions et de la chance. La durée moyenne de chaque séance journalière est d'environ 20 minutes.

Au début de l'expérience, les participants reçoivent une fiche des instructions, une feuille de décision, ainsi qu'un stylo[6]. L'expérience de risque se divise en deux volets. Sur la fiche d'expérience[7], les participants doivent effectuer 10 décisions qui consistent chacune à indiquer leur préférence entre deux

[5] Il est important de mentionner que la présence d'un effet de pairs lié à la productivité/diffusion/concurrence entre travailleurs n'affecterait pas les résultats si cet effet est similaire dans les deux contrats. Toutefois, si l'effet est différent d'un contrat à l'autre, par exemple si la compétition est plus forte dans le contrat à la pièce, la règle du choix du contrat peut être affectée. Cette question peut constituer une piste intéressante pour de futures recherches.

[6] Les participants sont informés qu'ils effectueront 10 décisions et que l'une de leurs décisions sera tirée au hasard et jouée pour déterminer le gain additionnel qui sera ajouté à leur prochain chèque de paie. Après avoir lu et expliqué les instructions à haute voix, les chercheurs ont autorisé les participants à poser les questions d'éclaircissement sur la procédure de l'expérience. Ensuite, chaque participant remplit sa fiche individuellement et en silence. L'équipe assure le suivi pendant l'expérience pour répondre à toute question des participants.

[7] voir table 1.10

loteries: une loterie "A" (moins risquée) qui paye soit 40$ ou 32$ et une loterie "B" (plus risquée) qui paye soit 77$ ou 2$ [8]. Ces valeurs sont fixes à travers les 10 décisions. La fiche de décision est présentée en annexe au tableu 1.10. À chacun de ces gains possibles est associé une probabilité qui varie selon la décision. Pour simplification, "A" sera référée comme la "loterie sûre" et "B", la "loterie risquée". À chaque décision, la probabilité de succès (ou échec) est identique pour les deux loteries. À la première décision, la probabilité de perdre au jeu (gagner 32$ si l'on choisit la loterie "A" et 2$ si l'on choisit "B") est fixée à 90%. Dans ce cas, seul un agent extrêmement risquophile préféra jouer la loterie "B", l'espérance de gain de la loterie sûre étant nettement plus élevée pour la loterie A que pour la loterie "B"[9]. En descendant sur la fiche de décision, la probabilité de succès augmente par 1/10 à chaque décision. En l'occurrence, la probabilité de succès atteint $\frac{1}{2}$ à la décision 5. À la dernière décision, tout participant rationnel choisira la loterie "B", car elle paye 77$ avec certitude tandis que la loterie "A" paye 40$. En descendant sur la fiche, l'espérance de gain de la loterie risquée augmente relativement à celle de la loterie sûre. Alors chaque participant choisira éventuellement la loterie B à l'une de ses décisions. Le "point de passage", qui définit le nombre de "choix-sûrs" que l'individu effectue avant de passer à la loterie risquée, permet de décrire ses préférences pour le risque. Par exemple, un agent neutre au risque choisit la loterie A à la décision 4 et la loterie B à la décision 5. Si l'agent effectue moins de 4 choix-sûrs, alors il est risquophile et s'il fait plus de 4 choix-sûrs, cela révèle un agent averse au risque. Après avoir fini de remplir la fiche de choix, les participants tirent avec remise deux boules d'un sac contenant 10 boules numérotées de 1 à 10 et indiscernables au toucher. Le premier tirage détermine la loterie qui sera jouée. Le second tirage détermine le résultat de la loterie et donc le gain du participant.

Une question importante lorsque l'on mesure les préférences pour le risque en utilisant des loteries est de savoir si la décision des participants est affectée ou non par le montant des mises. Par exemple, les participants pourraient considérer que les montants en jeu ne sont pas suffisamment élevés pour être pris au sérieux[10].

Afin de prendre cela en compte, le jour suivant la première expérience, nous avons répliqué sur les mêmes travailleurs l'expérience de mesure des préférences pour le risque précédente en doublant les gains des loteries, mais en gardant les frais de participation à 20$ pour cette expérience également[11]. Ce traitement sera noté "Loterie avec des Mises Élevées (LME)", le premier étant désigné "Loterie avec des Mises Faibles(LMF)" [12]. Dans ce nouveau traitement, la loterie "A" paye alors 80$ pour "succès" et 64$ pour "échec" tandis que la loterie "B" paye soit 154$ soit 4$. Le gain final du participant est obtenu suivant les mêmes procédés par des tirages aléatoires.

L'avantage de la méthode expérimentale de Holt and Laury (2002) est de permettre une estimation directe des préférences pour le risque à partir du choix des loteries. Toutefois, l'un des inconvénients majeur de cette méthode est sa complexité et la possibilité que les participants ne comprennent pas l'expérience Charness et al. (2013). Une façon de mesurer la compréhension de l'expérience est de

[8]le choix de ces montants pour la lotterie est identique à ceux de Bellemare and Shearer (2010) et Bellemare and Shearer (2013) qui ont effectué des expériences similaires.

[9]Voir le différentiel de gain espéré entre la loterie B et la loterie A E(B)-E(A) sur le tableau 1.10

[10]Les préférences pour le risque peuvent aussi se modifier avec la taille des mises (voir Holt and Laury (2002).)

[11] Un nouveau travailleur était présent ce jour et a participé à cette étape

[12]Une copie de chaque feuille se trouve en annexe

calculer le pourcentage de choix incohérents. Les choix du travailleur sont dits "incohérents" s'il a plus d'un point de passage de la loterie A à la loterie B ou si l'agent choisi la loterie A à toutes les décisions. Dans le cas de notre expérience, seulement 2 participants (soit 3,2%) ont effectué des choix incohérents durant l'expérience avec des mises faibles; ce qui constitue une proportion relativement faible, comparativement à d'autres études telles que Charness and Viceisza (2011) et Jacobson and Petrie (2009) qui ont trouvé respectivement 75% et 55% de choix incohérents.

1.5.2 La mesure de la préférence pour le risque

Le nombre de "choix-sûrs" effectués par l'agent économique dans l'expérience de risque de Holt and Laury (2002) peut être associé à une fonction d'utilité représentant les préférences de risque. En considérant une aversion relative au risque constante pour le gain x, les préférences peuvent se représenter par la fonction suivante: $U(x) = \frac{1}{\delta}x^{\delta}$ ou $x > 0$ et δ est le coefficient d'aversion relative au risque constant. Cette spécification traduit un comportement risquophile si $\delta < 1$, neutre au risque si $\delta = 1$ et d'aversion au risque si $\delta > 1$.

Le point de passage de la loterie "A" à la loterie "B" donne un intervalle de l'aversion au risque relatif estimé du participant. Le tableau (1.1) représente la classification des préférences pour le risque, dérivée de la fonction $U(x) = \frac{1}{\delta}x^{\delta}$ selon le nombre de choix-sûrs effectués par les agents. En illustration, étant donné les montants des loteries A et B, un choix de 4 loteries sûres suivi de 6 loteries risquées correspond à une aversion relative au risque constante comprise dans l'intervalle [0,85 ; 1,14].

Table 1.1: Classe de préférences pour le risque selon le nombre de choix-sûrs effectués.

Nombre de choix-sûrs	$U(x) = \frac{1}{\delta}x^{\delta}$	Classe de préférences pour le risque
0-1	$\delta \geq 1.95$	Risquophile extrême
2	$1.49 \leq \delta \leq 1.95$	Fortement risquophile
3	$1.14 \leq \delta \leq 1.49$	Risquophile
4	$0.85 \leq \delta \leq 1.14$	Neutre au risque
5	$0.59 \leq \delta \leq 0.85$	Légèrement averse au risque
6	$0.32 \leq \delta \leq 0.59$	Averse au risque
7	$0.029 \leq \delta \leq 0.32$	Fortement averse au risque
8	$-0.37 \leq \delta \leq 0.029$	Très fortement averse au risque
9-10	$\delta \leq -0.37$	Averse au risque extrême

Note: La colonne 1 est le nombre de choix-sûrs. La colonne 2 indique l'intervalle du coefficient constant de préférence pour le risque et la colonne 3 représente la classification correspondante.

1.5.3 L'expérience de choix de contrat

Si les préférences mesurées par l'expérience de risque expliquent le comportement des travailleurs de cette firme en présence de risque, elles devraient prédire le choix des planteurs entre un contrat à salaire fixe et un contrat à la pièce. Dans la suite de l'expérience de risque, une expérience de choix de contrat est alors effectuée sur les mêmes participants afin de tester si les résultats de l'expérience de risque sont compatibles avec les décisions d'auto-sélection entre les contrats. Les participants recoivent une compensation initiale forfaitaire de 20$ pour sa participation à cette expérience également. Cette expérience s'est déroulée du 30 mai au 03 juin 2016 regroupant 50 planteurs[13]. Dans ce deuxième volet, les participants doivent effectuer 12 décisions dont chacune consiste à choisir entre leur contrat à la pièce et un contrat à salaire fixe[14]. Le contrat à la pièce qui rémunère les travailleurs selon leur performance journalière est le mode de payement habituellement utilisé par la firme. Notez que le salaire du travailleur sous le contrat à la pièce est soumis au risque tandis que le contrat fixe assure un salaire certain. Sous un contrat à la pièce, les gains sont déterminés par la production et le taux à la pièce. Ces gains sont aléatoires, car la production évaluée en nombre d'arbres plantés dépend des conditions inobservables du terrain. De plus, au moment de l'expérience, le planteur est informé de la parcelle où il est assigné, mais il ne connaît pas le positionnement exact qui sera le sien sur cette parcelle. Pour effectuer son choix de contrat, le planteur compare donc son utilité espérée en contrat à la pièce au salaire fixe qui lui est proposé. Il choisit le contrat à la pièce si son gain espéré est plus grand. Dans le cas contraire, il choisit le contrat fixe.

A la fin de l'expérience, pour chaque travailleur, l'une de ces décisions a été tirée au hasard et les travailleurs ont été assignés au contrat qu'ils ont choisi à cette décision pour les deux prochains jours de travail. Sur la fiche de choix de contrat (présentée en annexe tableau 1.12), le contrat à taux fixe à la première décision est de 150$. Étant donné que le salaire moyen des travailleurs est de 417$ par jour quand ils sont payés à la pièce, seul un agent peu productif et/ou extrêmement averse au risque choisirait le contrat fixe à cette décision. Le contrat à taux fixe s'accroît de 50$ par jour pour les autres décisions pour atteindre le montant maximal de 650$ par jour à la douzième décision. Le contrat fixe devient alors de plus en plus attrayant. Un agent dont l'espérance de gain sous un contrat à la pièce est inférieure au montant maximal du payement fixe choisira éventuellement un contrat fixe à l'une de ses décisions.

Tout comme dans l'expérience de loteries, le contrat de chaque travailleur est déterminé par l'aléa. Chaque travailleur tire une boule d'un sac contenant 12 boules, énumérées de 1 à 12. Si le travailleur tire le numéro 4 et si le travailleur a choisi le contrat à la pièce pour la décision 4, il est rémunéré suivant son contrat à la pièce régulier. S'il a choisi le salaire fixe, il reçoit un salaire fixe journalier de 200$ les deux prochains jours de travail. Le nombre de contrats à la pièce choisis par l'agent avant de passer au contrat fixe décrit la valeur que l'individu accorde au contrat à la pièce en termes de contrat fixe.

1.6 Les données

Les données utilisées dans ce papier proviennent de deux sources: les données expérimentales et les données récoltées sur le terrain pendant la campagne de reboisement.

1.6.1 Les données expérimentales

Les expériences pour mesurer les préférences pour le risque

Les résultats de l'expérience de risque sont présentés dans le tableau 1.2. Tous les travailleurs présents ont participé à l'étude, ce qui exclut la possibilité d'un effet de sélection. Les statistiques descrip-

[13] Tous les travailleurs présents ce jour ont participé à l'expérience. Les participants reçoivent également une somme forfaitaire de 20$ pour leur participation.

[14] Voir fiche du choix de contrat Tableau 1.12

tives montrent que seulement 34,37% des travailleurs manifestent une aversion pour le risque lors de l'expérience avec des mises faibles. Cela suggère qu'une proportion importante, soit 65,63% des participants sont, soit risquophiles soit neutres au risque[15]. Lors du jeu avec des mises faibles, les gains obtenus par les participants se situent entre 22$ et 97$ avec une moyenne de 74.46$ et une variance de 25.45$ [16].

En doublant les mises des loteries, la proportion de planteurs averses au risque s'accroît d'environ 5 points. Les gains obtenus durant ce traitement se situent entre 24$ et 174$ avec une moyenne de 104.28$.

Table 1.2: Distribution des Choix de la loterie sûre.

		LME		LMF	
Choix sûrs	CRRA pour $U = \frac{1}{\delta}x^{\delta}$	Ensemble	Cohérent	Ensemble	Cohérent
0-1	$\delta \geq 1.94684$	1.56	1.56	4.76	4.92
2	$1.48657 \leq \delta \leq 1.94684$	1.56	1.56	6.35	6.56
3	$1.14263 \leq \delta \leq 1.48657$	17.19	17.19	25.40	26.23
4	$0.853637 \leq \delta \leq 1.14263$	40.63	40.63	42.86	40.98
5	$0.588544 \leq \delta \leq 0.853637$	65.63	65.63	61.90	60.66
6	$0.32382 \leq \delta \leq 0.588544$	85.94	85.94	85.71	85.25
7	$0.0294191 \leq \delta \leq 0.32382$	95.31	95.31	96.83	96.72
8	$-0.368391 \leq \delta \leq 0.0294191$	98.44	98.31	98.41	98.36
9-10	$\delta \leq -0.368391$	100	100	100	100
Échantillon		64	64	63	61

Le choix de contrat

La figure 1.1 présente la répartition du nombre de contrats à la pièce choisis. Le nombre moyen de contrats à la pièce choisis est de 7,4 soit un équivalent fixe moyen du contrat à la pièce se situant entre 400 et 450$. Dans cette expérience, un travailleur a choisi le contrat à la pièce pour toutes ses décisions. Cela signifie que la valeur que ce dernier accorde au contrat à la pièce est supérieure à la valeur maximale du salaire fixe de 650$ proposée à la dernière décision. La figure 1.2 présente la

[15]Ce résultat rejoint celui de Bellemare and Shearer (2010) qui trouvent une forte tolérance pour le risque au niveau des planteurs d'arbres, mais apparaît plus élevé que celui de Holt and Laury (2002) pour un groupe d'étudiants.

[16]En incluant les frais de participation

distribution cumulée du nombre de choix de contrats à la pièce par décision en fonction de la valeur du contrat à salaire fixe.

Figure 1.1: Répartition du nombre de contrats à la pièce choisis.

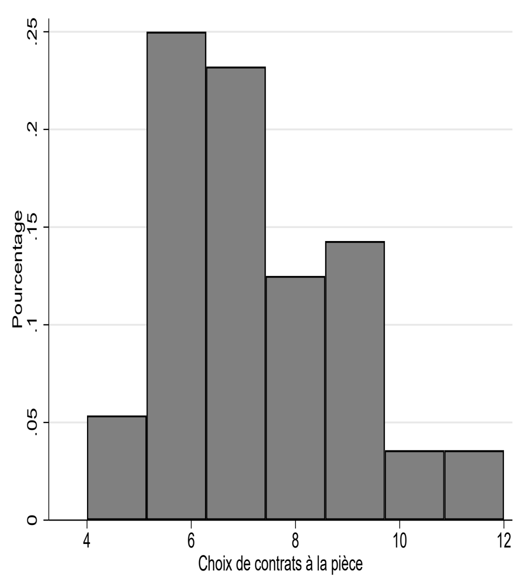
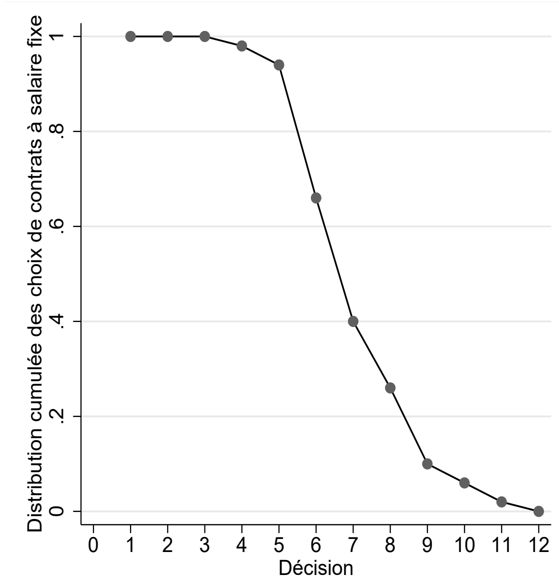


Figure 1.2: Distribution cumulée des choix de contrats à la pièce par décision.



Les statistiques de l'expérience de choix de contrat sont présentées dans le tableau 1.3. Il apparaît qu'une majeure partie des travailleurs, soit 76% acceptent les contrats fixes supérieurs ou égaux à 400$ par jour[17].

Le nombre de contrats à la pièce choisis permet d'identifier empiriquement l'intervalle de l'équivalent certain qui correspond à la valeur donnée au contrat à la pièce par chaque travailleur. Par exemple, un travailleur qui choisit sept contrats à la pièce avant de passer au contrat fixe refuse d'être payé à 400$ par jour, mais est prêt à travailler sous un contrat journalier de 450$. La valeur W^* qu'il accorde au contrat à la pièce en termes de contrat fixe se situe donc dans l'intervalle]400, 450].

1.6.2 Les données de reboisement

Ces données regroupent les informations sur 4106 observations[18] récoltées dans une firme de reboisement en Colombie- Britannique pendant la saison de plantation d'avril à juin 2016. Les données contiennent les informations sur la production journalière de chaque planteur (le nombre d'arbres plantés), la parcelle sur laquelle l'agent a travaillé, le taux à la pièce et les caractéristiques individuelles: l'âge, le sexe, la tenure, l'expérience et le niveau d'éducation.

Les statistiques descriptives des données liées au travail sont présentées dans le tableau 1.4. À la lumière de ces résultats, il ressort que les employés ont travaillé en moyenne 42 jours pendant cette campagne. Les salaires journaliers se situent entre 101,25$ et 1065,05$ avec une moyenne 420,91$.

[17] Ce contrat paye juste en-dessous de la moyenne des gains réels pendant le reboisement soit 420,9$ par jour.

[18] Nous retenons seulement les observations ou le gain est au moins supérieur au salaire minimum garanti de 83.6$ par jour en Colombie Britannique. Cela exclut les planteurs ayant travaillé une partie de la journée.

Table 1.3: Choix de contrats

Nombre de contrats à la pièce	Intervalle de l'équivalent certain	Distribution cumulée
0-4	$W^* \leq 300$	2,00
5	$200 < W^* \leq 350$	6,00
6	$350 < W^* \leq 400$	34,00
7	$400 < W^* \leq 450$	60,00
8	$450 < W^* \leq 500$	74,00
9	$500 < W^* \leq 550$	90,00
10	$550 < W^* \leq 600$	94,00
11	$600 < W^* \leq 650$	98,00
12	$W^* > 650$	100

Note: La première colonne du tableau représente le nombre de contrats à la pièce choisis par le participant avant de passer au contrat fixe. La deuxième colonne définit les bornes du salaire fixe équivalent au contrat à la pièce. La dernière colonne représente la distribution cumulée correspondante.

Les travailleurs plantent en moyenne 2005 arbres par jour avec un écart-type de 622$. L'analyse des données par type de contrat de travail montre que la productivité s'étend entre 330 et 4915 arbres plantés sous le taux à la pièce régulier. De plus, la variabilité de la productivité est plus forte en contrat à la pièce (écart-type de 620 plants) qu'en contrat fixe (erreur standard de 389 plants). Ceci traduit bien l'ampleur du risque financier associé au contrat à la pièce. Les gains des travailleurs sous contrat à la pièce sont moins élevés 417,04$ contre 527,77$ en contrat fixe. Cela s'explique par le fait que le travailleur choisit le contrat à taux fixe lorsque les salaires sont élevés, ce qui génère une rente. En effet, si les travailleurs en contrat fixe étaient payés à un salaire à la performance régulier, ils toucheraient un salaire moyen moins élevé soit 310,52$ par jour.

Table 1.4: Synthèse des statistiques sur la productivité et les gains

	Obs-jour	Moyenne	Stand. dev.	Min	Max
Ensemble					
Gains/jour	4106	420.9181	115.1793	101.25	1065.05
Arbres plantés	4106	2004.712	621.5831	330	4915
Jours de travail par travailleur.	2354	41.81478	5.900335	2	50
Contrat à la pièce					
Taux à la pièce	3978	.2130901	.0366968	.17	.38
Arbres plantés	3978	2005.192	619.6768	330	4915
Gains/jour en CP	3978	417.0392	112.573	101.25	933.85
Contrat fixe(CF)					
Arbres plantés	36	1605.833	388.9005	585	2450
Gains/jour en CF	36	527.7778	74.10846	400	650
Contrefactuel					
Contrefactuel des gains en CF	36	310.526	73.1584	105.3	465.5

Note: Les rangées de ligne 1 à 4 définissent les statistiques descriptives respectivement pour l'ensemble des planteurs, les planteurs travaillant sous un contrat à la pièce régulier, les contrats à la pièce augmentés et les travailleurs sous contrat fixe. La dernière rangée de ligne représente un contre-factuel donnant les gains que les travailleurs en contrat fixe auraient gagnés s'ils étaient payés à la pièce.

Le tableau 1.5 représente les statistiques descriptives des caractéristiques individuelles. Ce tableau montre que les travailleurs de cette entreprise constituent un groupe hétérogène. L'âge des travailleurs se situe entre 20 et 47 ans et 45,76% d'entre eux sont des femmes. Du côté de l'éducation qui représente le diplôme le plus élevé acquis par le travailleur, nous observons que 66,1% des employés possèdent un diplôme universitaire. [19] Par ailleurs, cette cohorte comprend des travailleurs d'expérience comptabilisant jusqu'à 28 ans d'expérience, mais aussi des planteurs avec seulement un an d'expérience. La variable "tenure" représente le nombre d'années d'ancienneté du travailleur dans cette entreprise. L'ancienneté moyenne parmi ces travailleurs est de 4,4 ans. Le travailleur le plus ancien dans cette entreprise totalise 26 ans et le moins ancien a un an dans la firme.

Table 1.5: Caractéristiques individuelles des travailleurs

Variable	Obs	Moyenne	Dev. Std.	Min	Max
Age	59	29,13559	4,542738	20	47
Experience	59	8,983051	4,984429	1	28
Tenure	59	4,372881	4,409622	1	26
Genre					
Femmes	59	0,4576271	0,5024778	0	1
Hommes	59	0,5423729	0,5024778	0	1
Niveau d'éducation					
Secondaire	59	0,0508475	0,2215719	0	1
Début Post-sec	59	0,0169492	0,1301889	0	1
Post-secondaire	59	0,0169492	0,1301889	0	1
Début universitaire	59	0,2542373	0,4391693	0	1
Diplôme Universitaire	59	0,6610169	0,4774274	0	1

1.7 Le modèle théorique

Le cadre théorique se fonde sur les travaux de Paarsch and Shearer (2000), Shearer (2004) et Bellemare and Shearer (2013). Considérons le cadre conceptuel d'un modèle principal-agent à une période comportant une firme qui engage plusieurs travailleurs i={1....N}. Les travailleurs sont hétérogènes en termes de préférences pour le risque et de capacités productives. La firme offre deux types de contrats: le contrat à la pièce qui paye selon la performance et un contrat à salaire fixe qui paye une rémunération par unité de temps de travail indépendamment de la production de l'employé.

1.7.1 La technologie

Supposons que les terrains à planter sont répartis en parcelles j={1,.....J} et que les travailleurs sont affectés chaque matin dans ces parcelles de façon aléatoire. La productivité du travailleur i dans la parcelle j noté $Y_{ij} > 0$ dépend de l'effort fourni par le travailleur $e_{ij} \in E \equiv [0, \bar{E}]$, mais aussi des conditions naturelles du terrain $s_j \in S$ qui sont aléatoires (pluviométrie, température, état du sol...). La technologie de production peut donc s'écrire:

[19] Ce qui se justifie par le fait que l'activité de plantation d'arbres regroupe en majorité des étudiants pendant la saison estivale.

$$Y_{ij} = e_{ij} s_j \tag{1.1}$$

Le choc de productivité s_j capture l'ensemble des perturbations du modèle indépendantes du travailleur et qui peuvent affecter sa production. Son logarithme naturel $ln(s_j)$ suit une loi normale de moyenne μ_j et de variance σ^2.

1.7.2 Les préférences

Les préférences du travailleur sur le salaire et l'effort sont représentés par une fonction d'utilité CRRA séparable, continue et deux fois différentiable. Étant donné que l'effort est inobservable, le salaire à la pièce w_{ij} est indexé sur le niveau de production observée Y_{ij} et le taux à la pièce r_j:

$$w_{ij} = r_j Y_{ij} \tag{1.2}$$

Par contre, dans le contrat à salaire fixe, le travailleur reçoit un paiement w^f indépendant de sa production mais accepte de fournir un minimum d'effort. Un contrôle du travailleur par les superviseurs constitue une pression suffisante pour inciter le travailleur à fournir un minimum d'effort sous le contrat à taux fixe noté e^{fw}. L'utilité s'écrit:

$$U(w_{ij}, e_{ij}) = \begin{cases} \frac{1}{\delta_i}[w_{ij} - C(e_{ij})]^{\delta_i} & \text{if } w_{ij} > C(e_{ij}) \\ -\infty & \text{sinon,} \end{cases} \tag{1.3}$$

où $C(e_{ij}) = \frac{1}{\eta} k_i e_{ij}^{\eta}$ est le coût de l'effort représentant une désutilité pour le travailleur; δ_i l'aversion relative au risque constante et k_i est un indice des capacités productives intrinsèques du travailleur. Une hausse de k_i révèle un agent aux capacités productives plus faibles car le coût marginal de l'effort s'accroît. Le paramètre η défini la courbure de la fonction de coût. Cette dernière est définie telle que $C'(.) > 0$, $C''(.) > 0$, C(0)=0 et C"(0)=0. Cette spécification assure l'existence d'un optimum fini pour le choix de l'effort.

1.7.3 Les étapes du contrat

La séquence de mise en œuvre du contrat se présente comme suit:

1. La nature choisit (μ_j, σ^2) tel que le choc de productivité moyen μ_j varie selon la parcelle.
2. La firme observe σ^2, mais prédit $\tilde{\mu}_j$ tel que $\mu_j = \tilde{\mu}_j + \epsilon_j$ avec $\epsilon_j \sim N(0, \sigma^2_\epsilon)$.
3. Étant données les valeurs de σ^2 et $\tilde{\mu}_j$, la firme propose au travailleur deux types de paiement:

$$\Omega = \begin{cases} r & \text{le contrat à la pièce} \\ w_j^f & \text{le contrat à salaire fixe} \end{cases} \tag{1.4}$$

4. Le travailleur observe $\tilde{\mu}_j$, σ^2 et Ω et choisit soit le contrat à la pièce soit le contrat à salaire fixe
5. Le travailleur observe s_{ij} et décide de son niveau d'effort e_{ij}.
6. La firme observe la production Y_{ij} et paye les salaires convenus par le contrat.

1.7.4 Le contrat à la pièce

Après avoir choisi le contrat à la pièce et observé l'état de la nature, le programme de l'agent consiste à choisir l'effort optimal e_{ij}^p qui maximise son utilité étant donné la réalisation s_j. En résolvant le problème de l'agent, l'effort optimal s'écrit:

$$e_{ij}^p = \left[\frac{r_j s_j}{k_i}\right]^{\gamma} \text{ avec } \gamma = \frac{1}{\eta - 1}. \tag{1.5}$$

Observons que l'effort optimal en (1.5) est indépendant du risque car l'effort est choisi après avoir observé S_j. La condition de second ordre: $\frac{\partial^2 U}{\partial e_{ij}^2} < 0$[20] assure l'existence d'un effort optimal permettant de maximiser l'utilité du travailleur.

Étant donné r_j, les conditions du terrain s_j et l'effort optimal e_{ij}^P, le salaire est :

$$w(s_{ij}) = \frac{r_j^{\gamma+1}}{k_i^{\gamma}} s_j^{\gamma+1} \tag{1.6}$$

L'agent supporte une désutilité de l'effort $C(e_{ij})$ qui s'écrit :

$$C(e_{ij}) = \frac{\gamma}{\gamma+1} \frac{r_j^{\gamma+1}}{k_i^{\gamma}} s_j^{\gamma+1} \tag{1.7}$$

L'utilité indirecte de chaque travailleur V_{ij}^p / s_{ij} est:

$$V^p(s_j) = \frac{1}{\delta_i} \frac{r_j^{\delta_i(\gamma+1)}}{(\gamma+1)^{\delta_i} k_i^{\delta_i \gamma}} s_j^{\delta_i(\gamma+1)} \tag{1.8}$$

L'utilisation des propriétés des fonctions log-normales permet de dériver à partir des équations (1.6), (1.7) et (1.8) , les valeurs espérées du salaire (1.9), du coût de l'effort induit (1.10) et de l'utilité indirecte (1.11):

$$E\left[w_{ij}^p(k_i, k_h, \mu_j/\sigma_j)\right] = \frac{r_j^{\gamma+1}}{k_i^{\gamma}} \exp^{(\gamma+1)\mu_j + \frac{1}{2}(\gamma+1)^2\sigma^2} \tag{1.9}$$

$$E\left[c(e^{*p})\right] = \frac{\gamma}{\gamma+1} \frac{r_j^{\gamma+1}}{k_i^{\gamma}} \exp^{(\gamma+1)\mu_j + \frac{1}{2}(\gamma+1)^2\sigma^2} \tag{1.10}$$

$$E\left[V_{ij}^p/\mu_j)\right] = \frac{1}{\delta_i} \frac{r_j^{\delta_i(\gamma+1)}}{(\gamma+1)^{\delta_i} k_i^{\delta_i\gamma}} \exp^{\delta_i(\gamma+1)\mu_j + \frac{1}{2}(\gamma+1)^2 {\delta_i}^2 \sigma^2} \tag{1.11}$$

$$E\left[V_{ij}^p/\tilde{\mu}_j)\right] = \frac{1}{\delta_i} \frac{r_j^{\delta_i(\gamma+1)}}{(\gamma+1)^{\delta_i} k_i^{\delta_i\gamma}} \exp^{\delta_i(\gamma+1)\tilde{\mu}_j + \frac{1}{2}(\gamma+1)^2 {\delta_i}^2 \sigma^2} E\left(\exp^{\delta_i(\gamma+1)\epsilon_j}\right) \tag{1.12}$$

[20] $\frac{\partial^2 U}{\partial E_{ij}^2} = -\frac{1}{\gamma} k_i^{\gamma} (r_j s_j)^{1-\gamma}$

Étant donnée que $\epsilon_j \sim N(0, \sigma_\epsilon^2)$, en appliquant les propriétés de la loi normale, l'utilité indirecte devient:

$$E\left[V_{ij}^p/\tilde{\mu}_j)\right] = \frac{1}{\delta_i} \frac{r_j^{\delta_i(\gamma+1)}}{(\gamma+1)^{\delta_i} k_i^{\delta_i \gamma}} \exp^{\delta_i(\gamma+1)\tilde{\mu}_j + \frac{1}{2}(\gamma+1)^2 \delta_i{}^2(\sigma^2+\sigma_\epsilon^2)} \tag{1.13}$$

La détermination du taux à la pièce par la firme doit satisfaire les contraintes de participation de l'agent. Posons $\bar{w}$ le salaire alternatif du marché et h le travailleur marginal défini comme indifférent entre le contrat à la pièce et cette alternative. La firme choisit le taux à la pièce r_j tel que:

$$\frac{1}{\delta_h} \frac{r_j^{\delta_h(\gamma+1)}}{(\gamma+1)^{\delta_h} k_h^{\delta_h \gamma}} \exp^{(\delta_h(\gamma+1)\tilde{\mu}_j + \frac{1}{2}(\gamma+1)^2 \delta_h{}^2(\sigma^2+\sigma_\epsilon^2))} = \frac{1}{\delta_h} \bar{w}^{\delta_h} \tag{1.14}$$

$$r_j^{\gamma+1} = \bar{w}(\gamma+1) k_h^\gamma \exp^{-\left[(\gamma+1)\tilde{\mu}_j + \frac{1}{2}\delta_h(\gamma+1)^2(\sigma^2+\sigma_\epsilon^2)\right]} \tag{1.15}$$

En substituant $r_j^{\gamma+1}$ de (1.15) dans (1.13), l'utilité espérée du contrat à la pièce se réécrit :

$$E\left[V_{ij}^p(\delta_i, k_i)\right] = \frac{1}{\delta_i} \left[\bar{w} \left(\frac{k_h}{k_i}\right)^\gamma\right]^{\delta_i} \exp^{\frac{1}{2}\delta_i(\gamma+1)^2(\delta_i - \delta_h)(\sigma^2+\sigma_\epsilon^2)} \tag{1.16}$$

L'expression (1.16) donne l'utilité indirecte que le travailleur peut espérer atteindre étant données sa préférence pour le risque et ses capacités productives s'il décide de travailler sous le contrat à la pièce. Cette utilité ne dépend pas de la parcelle j.

1.7.5 Le contrat à salaire fixe

Dans le contrat à salaire fixe, l'agent reçoit un paiement par unité de temps de travail indépendant de sa production. Il doit fournir un minimum d'effort. Nous supposons que l'effort dans le contrat à salaire fixe est proportionnel à l'effort optimal dans le contrat à la pièce[21]. Cette hypothèse s'écrit alors: $e_{ij}^{fw} = \psi_i E(e^{*p})$ [22]

En remplaçant la valeur espérée de l'équation (1.5) dans l'hypothèse 1 , l'effort du travailleur sous le contrat à salaire fixe peut alors s'écrire :

$$E\left[e_{ij}^{fw}/\mu_j\right] = \psi_i \frac{r_j^\gamma \exp^{\gamma\mu_j + \frac{1}{2}\gamma^2\sigma^2}}{k_i^\gamma} \tag{1.17}$$

Étant donné que la firme et le travailleur observent $\tilde{\mu}_j$ plutôt que μ_j alors:

$$E\left[e_{ij}^{fw}/\tilde{\mu}_j\right] = \psi_i \frac{r_j^\gamma \exp^{\gamma\tilde{\mu}_j + \frac{1}{2}\gamma^2(\sigma^2+\sigma_\epsilon^2)}}{k_i^\gamma} \tag{1.18}$$

Dans ce cas, l'utilité de l'agent dans le contrat à salaire fixe est:

$$V_{ij}^f = \frac{1}{\delta_i}\left[W_f - c(e_{ij}^{fw})\right]^{\delta_i} \tag{1.19}$$

De (1.19), on dérive l'utilité indirecte sous le contrat à salaire fixe:

$$E\left[V_{ij}^f(\delta_i, k_i, W_f)\right] = \frac{1}{\delta_i}\left[W_f - \gamma\psi_i^{\frac{\gamma+1}{\gamma}} \bar{w} \left(\frac{k_h}{k_i}\right)^\gamma \exp^{\frac{1}{2}\gamma(\gamma+1)(\sigma^2+\sigma_\epsilon^2) - \frac{1}{2}\delta_h(\gamma+1)^2(\sigma^2+\sigma_\epsilon^2)}\right]^{\delta_i} \tag{1.20}$$

[21]Cette hypothèse semble être cohérente avec la politique de l'entreprise.

[22]Avec $(1-\psi_i)$ la baisse de l'effort de l'agent du contrat à la pièce au contrat à salaire fixe.

L'équation (1.20) donne l'utilité indirecte que le travailleur peut espérer atteindre étant données ses préférences pour le risque, ses capacités productives et la fraction d'effort qu'il fournira s'il choisit de travailler sous le contrat à salaire fixe W_f qui lui est proposé par la firme. Cette utilité est indépendante de la parcelle.

1.7.6 Le choix du contrat et les prédictions du modèle

Cette section aborde la règle de sélection du contrat préféré par l'agent i en fonction de ses caractéristiques individuelles et de son programme de maximisation.

Le choix du contrat

Considérons un modèle avec un agent, une période et deux alternatives A et B en l'occurrence un contrat à salaire fixe et le contrat à la pièce. L'agent fait face à un choix binaire dont la règle de décision consiste à choisir l'alternative qui procure la meilleure utilité espérée. Le choix de l'agent dépend du vecteur de ses caractéristiques socio-économiques $z_i = (\delta_i, k_i)$ avec δ_i la préférence pour le risque et k_i l'indice des capacités productives. Ce choix dépend également des particularités de chaque contrat notamment le niveau de risque intrinsèque à chaque contrat. Le programme de l'agent se présente comme suit:

$$Max \left\{ E\left[V_{ij}^{pr}(\delta_i, k_i)\right], E\left[V_{ij}^{fw}(W_f, \delta_i, k_i)\right] \right\} \tag{1.21}$$

Un travailleur ayant des caractéristiques socio-économiques z_i choisit l'alternative "contrat à la pièce" si :

$$E\left[V_{ij}^{pr}(\delta_i, k_i)\right] > E\left[V_{ij}^{fw}(W_f, \delta_i, k_i)\right] \tag{1.22}$$

tandis que le contrat à salaire fixe sera préféré si:

$$E\left[V_{ij}^{pr}(\delta_i, k_i)\right] < E\left[V_{ij}^{fw}(W_f, \delta_i, k_i)\right] \tag{1.23}$$

Le travailleur est indifférent si les deux contrats procurent la même utilité. Pour chaque contrat à la pièce, il existe un équivalent certain défini par le contrat à salaire fixe qui assurerait à l'agent le même niveau d'utilité que le contrat à la pièce étant donné ses caractéristiques individuelles $z_i = (\delta_i, k_i)$. Notons W_f^* l'équivalent certain du contrat à la pièce. W_f^* solutionne l'équation:

$$E\left[V^{pr}(\delta_i, k_i)\right] = E\left[V^{fw}(W_f^*, \delta_i, k_i)\right] \tag{1.24}$$

soit:

$$\frac{1}{\delta_i}\left[\bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma} \exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-\delta_h)(\sigma^2+\sigma_\epsilon^2)}\right]^{\delta_i} = \frac{1}{\delta_i}\left[W_f^* - \gamma\psi_i^{\frac{\gamma+1}{\gamma}}\bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma} \exp^{\frac{1}{2}\gamma(\gamma+1)(\sigma^2+\sigma_\epsilon^2)-\frac{1}{2}(\gamma+1)^2\delta_h(\sigma^2+\sigma_\epsilon^2)}\right]^{\delta_i} \tag{1.25}$$

En résolvant (1.25), l'équivalent certain du contrat à la pièce s'écrit:

$$W_f^*(k_i, \delta_i) = \bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma} \exp^{\frac{1}{2}(\gamma+1)^2(\sigma^2+\sigma_\epsilon^2)(\delta_i-\delta_h)} + \gamma\psi_i^{\frac{\gamma+1}{\gamma}}\bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma} \exp^{\frac{1}{2}\gamma(\gamma+1)(\sigma^2+\sigma_\epsilon^2)-\frac{1}{2}(\gamma+1)^2\delta_h(\sigma^2+\sigma_\epsilon^2)} \tag{1.26}$$

En regroupant les termes de l'équation (1.26), on obtient:

$$W_f^*(k_i, \delta_i) = \bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma}\left[\exp^{\frac{1}{2}(\gamma+1)^2(\sigma^2+\sigma_\epsilon^2)(\delta_i-\delta_h)} + \gamma\psi_i^{\frac{\gamma+1}{\gamma}} \exp^{\frac{1}{2}\gamma(\gamma+1)(\sigma^2+\sigma_\epsilon^2)-\frac{1}{2}\delta_h(\gamma+1)^2(\sigma^2+\sigma_\epsilon^2)}\right] \tag{1.27}$$

L'équation (1.27) donne la valeur du contrat à salaire fixe qui procure au travailleur i la même utilité indirecte espérée que le contrat à la pièce. L'inspection de l'équation (1.26) montre que l'équivalent certain est constitué d'une première partie qui est la rente obtenue par le travailleur dans un contrat à la pièce et d'une seconde partie représentant la compensation du coût de l'effort du travailleur dans le contrat à salaire fixe.

La règle de décision est alors telle que le contrat à taux fixe est choisi si:

$$W_f > W_f^*(k_i, \delta_i) \tag{1.28}$$

et le contrat à la pièce est préféré si:

$$W_f < W_f^*(k_i, \delta_i) \tag{1.29}$$

Les prédictions testables

Toutes choses égales par ailleurs, l'équation (1.29) décrit l'ensemble de contrats à la pièce préférés au contrat fixe. Le gain marginal $G^*(k_i, \delta_i)$ du contrat à la pièce par rapport au contrat fixe est une variable latente décrivant l'écart entre les deux utilités espérées:

$$G^*(k_i, \delta_i) = W_f^*(k_i, \delta_i) - W_f \tag{1.30}$$

$G^*(k_i, \delta_i)$ représente le bénéfice pour le travailleur de choisir un contrat à la pièce plutôt qu'un contrat fixe et constitue la variable de choix du travailleur.

La règle de décision, présenté dans le figure (1.3), est telle que:

- Le travailleur choisira le contrat à la pièce si $G^*(k_i, \delta_i) > 0$, car dans cette zone, le contrat fixe proposé est plus faible que l'équivalent fixe du contrat à la pièce.

- Le travailleur choisira le contrat à salaire fixe si $G^*(k_i, \delta_i) < 0$, car dans cette zone, le contrat fixe proposé est plus élevé que l'équivalent fixe du contrat à la pièce.

Prédiction 1: La probabilité de choisir un contrat à la pièce s'accroît avec la préférence pour le risque, car la rente marginale du contrat en fonction de la préférence pour le risque δ_i est croissante.

$$\frac{\partial G^*(k_i, \delta_i)}{\partial \delta_i} > 0 \tag{1.31}$$

Sur le graphique (1.4), on observe qu'une hausse de δ_i entraîne un déplacement de la courbe de $G^*(W_f, k_i, \delta_i)$ vers la droite (hausse du gain marginal), ce qui implique un rallongement la zone où le travailleur choisi le contrat à la pièce au détriment du contrat à salaire fixe.

Prédiction 2: La probabilité de choisir un contrat à la pièce s'accroît avec la productivité, car le gain marginal du contrat à la pièce augmente avec les capacités productives du travailleur.

$$\frac{\partial G^*(k_i, \delta_i)}{\partial k_i} < 0 \tag{1.32}$$

Le graphique (1.5) montre qu'une augmentation de k_i (baisse de l'habileté) se traduit par un déplacement de la courbe de $G^*(W_f, k_i, \delta_i)$ vers la gauche ce qui augmente la zone où le contrat à salaire fixe est choisi.

Figure 1.3: Règle de choix du contrat

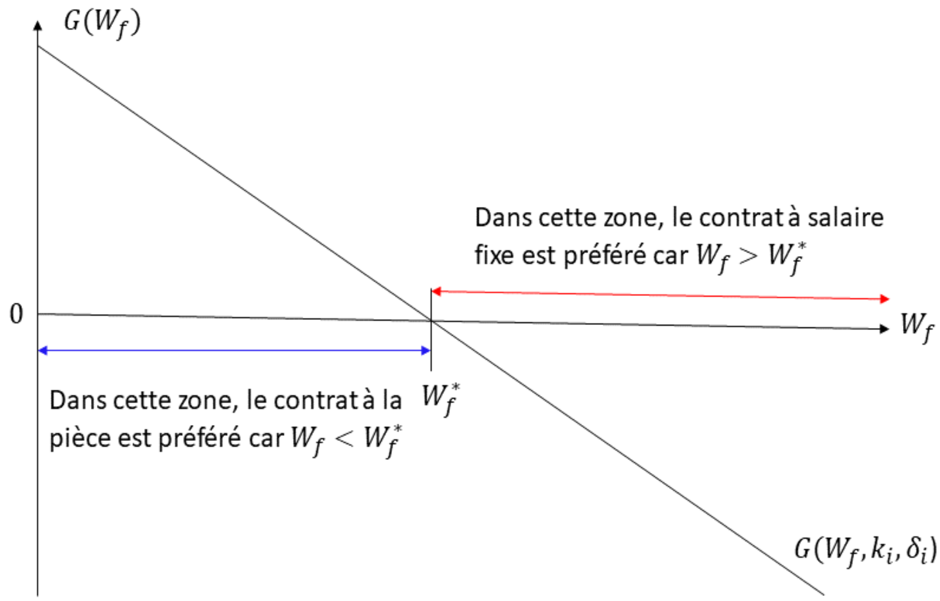


Figure 1.4: Hausse de δ_i ($\delta_i^b > \delta_i^a$)

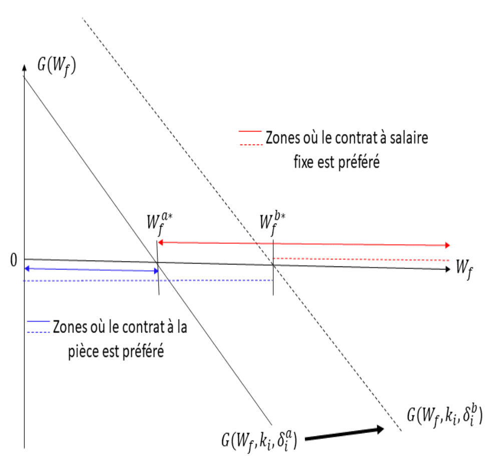
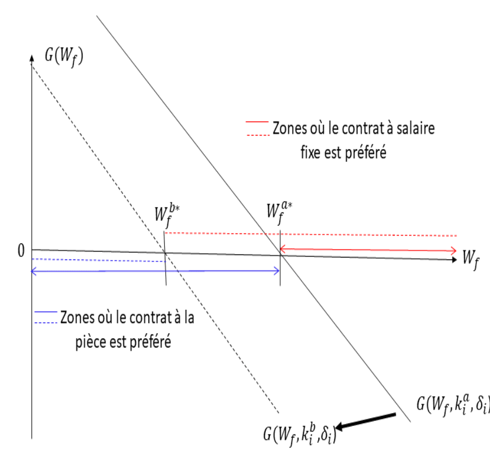


Figure 1.5: Hausse de k_i ($k_i^b > k_i^a$)



En combinant les données de l'expérience de risque et celles de l'expérience du choix de contrat, la suite de l'étude analyse la validité empirique de ces prédictions théoriques.

1.8 L'analyse de la validité interne des mesures expérimentales de préférence pour le risque

Cette section analyse la robustesse des mesures expérimentales de préférence pour le risque à des changements internes à l'expérience tel que la taille des mises utilisées dans les loteries.

1.8.1 Les déterminants de la préférence pour le risque

Les résultats de la section précédente suggèrent que la mesure de préférence pour le risque varie d'un traitement à l'autre. Dans cette section, nous évaluons l'impact des caractéristiques individuelles sur chacune de ces mesures en utilisant un modèle de Poisson. Soit N le nombre de participants et $cs_i \in CS = [0, 10]$ le nombre de fois où l'agent choisit la loterie-sûre. Le modèle de poisson est adéquat pour ces choix discrets. La moyenne conditionnelle s'écrit:

$$\lambda_i = exp(z_i'\beta) \tag{1.33}$$

où β représente le vecteur des paramètres et z_i le vecteur des variables explicatives incluant l'âge, le sexe, l'expérience, la tenure et l'éducation. La préférence pour le risque est une variable propre à l'individu et à ce titre, elle dépend de ses caractéristiques personnelles. La fonction de vraisemblance s'écrit:

$$Pr(CS = cs_i|z_i) = \frac{e^{-\lambda_i(z_i)}\lambda_i(z_i)^{s_i}}{S_i!} \tag{1.34}$$

On dérive les estimateurs de β en maximisant la log-vraisemblance:

$$LogL = \frac{1}{N}\sum_{i=1}^{N} log[Pr(CS = cs_i|z_i)] \tag{1.35}$$

Le tableau (1.6) représente les coefficients estimés du modèle de Poisson. Les résultats montrent que le nombre de choix-sûrs est négativement corrélé au niveau d'éducation du travailleur. Dans la mesure où le nombre de choix-sûrs capte l'aversion au risque, ceci suggère que les travailleurs les plus éduqués sont moins averses au risque. Par ailleurs, la tenure et l'expérience sont négativement corrélées avec l'aversion au risque mesurée seulement dans l'expérience LME. Ceci suggère que les travailleurs les plus expérimentés et ceux ayant passé plus d'années dans la firme sont moins averses au risque. La richesse cumulée par le travailleur n'a aucun effet sur sa préférence pour le risque dans les deux traitements. Ceci peut être une justification pour l'utilisation de la fonction d'aversion relative au risque constante pour décrire les préférences des travailleurs.

1.8.2 L'effet d'échelle

Une problématique fondamentale de l'évaluation expérimentale des préférences pour le risque est de savoir si oui ou non le comportement des agents en présence de risque est modifié lorsque les mises en jeu de chaque loterie sont plus élevées? Les résultats de Holt and Laury (2002) suggèrent que l'aversion au risque augmente lorsque la taille des mises s'accroît. Cependant, d'autres travaux empiriques

Table 1.6: Les déterminants de la préférence pour le risque.

	La variable dépendante est le nombre de choix-sûrs	
	LMF	LME
Gains précédents/100	-0,066	-0,001
	(0,141)	(0,133)
Age	-0,187*	0,087
	(0,099)	(0,106)
Age^2	0,003*	-0,001
	(0,099)	(0,002)
1 si femme	0,064	0,063
	(0,096)	(0,068)
Expérience	0,022	-0,024*
	(0,027)	(0,013)
Tenure	-0,063	-0,066***
	(0,04)	(0,013)
Éducation		
Post-secondaire	-0,219	0,012
	(0,165)	(0,161)
Debut universitaire	-0,652***	-0,24***
	(0,189)	(0,091)
Diplôme Universitaire	-0,584***	-0,315***
	(0,204)	(0,095)
Constante	4,91 ***	0,819
	(1,578)	(0,467)
Pseudo R2	0,08	0,02

Robust standard errors in parentheses

*** p<0,01, ** p<0,05, * p<0,1

Note: Les colonnes 2 et 3 présentent respectivement les coefficients de la régression du nombre de choix sûrs obtenu dans le traitement LME et dans le traitement LMF sur les facteurs individuels. Les erreurs standards robustes sont entre parenthèses.

n'observent pas de modification du comportement des agents (Bellemare and Shearer, 2010). Afin d'étudier cette question, nous testons d'abord si le nombre de choix-sûrs[23] est affecté par la taille des mises en utilisant le test de rang signé de Wilcoxon. L'hypothèse nulle du test est que la distribution du nombre de choix-sûrs est identique dans les deux traitements[24]. La statistique calculée du test est Z=-0.80 donnant une p-value de 0.424. Ce qui implique que l'hypothèse nulle n'est pas rejetée. Il n'y a donc pas de différences significatives dans les distributions des choix-sûrs entre ces deux traitements. On conclut donc que la distribution des préférences pour le risque est la même entre LMF et LME[25]. La figure (1.6) représente la distribution du nombre de choix-sûrs dans les deux traitements. La

[23] Rappelons que le nombre de choix-sûrs est le nombre de fois où la loterie A a été choisie

[24]Le principe du test consiste d'abord à calculer et ranger par ordre croissant les écarts en valeur absolue non nuls pour chaque paire. On note T la somme des rangs des écarts positifs. Cette somme est une des valeurs de la variable d'échantillonnage pour ce test. La table de Wilcoxon fournit la limite inférieure. Si T est plus élevé que sa borne théorique, on rejette l'hypothèse d'égalité des deux réponses.

[25]Le test de Khi-deux et le test de Kolmogorov-Smirnov ne rejettent pas non plus l'hypothèse d'égalité des deux distributions.

Figure 1.6: Répartition du nombre de choix-sûrs pour les deux traitements.

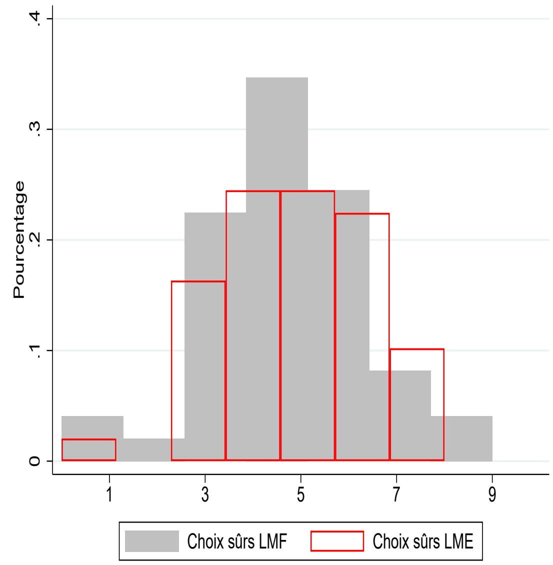


Figure 1.7: Distribution cumulée des choix-sûrs par décision pour les deux traitements.

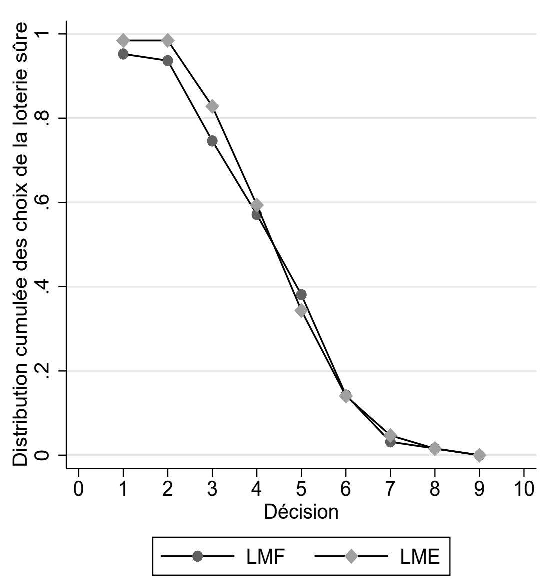


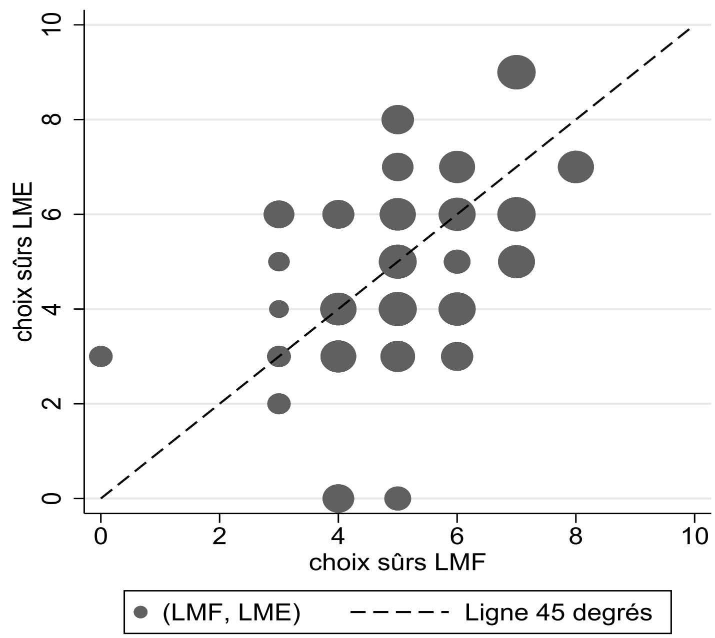
figure (1.7) représente la proportion de participants ayant choisi la loterie sûre par décision dans chaque traitement. Ces distributions apparaissent très similaires. Ce qui confirme les résultats du test de Wilcoxon que la distribution des préférences pour le risque ne change pas avec le traitement. Toutefois, ce résultat n'implique pas nécessairement une égalité des préférences capturées par ces deux mesures ou que les agents ne changent pas leur décision en fonction du traitement.[26]

La figure (1.8) montre le nombre des choix-sûrs par individu dans l'espace des deux traitements LME et LMF. Chaque point représente une combinaison de choix-sûrs effectués par un ou plusieurs participants dans les deux traitements LME et LMF. La taille de chaque point est proportionnelle à la fréquence d'individus ayant effectué la même combinaison de choix durant les deux traitements. La ligne de 45 degrés trace la fonction sur laquelle le nombre de choix-sûrs dans les deux traitements est égal. Si les planteurs ne changent pas leurs décisions, tous les points devraient s'étendre sur cette ligne. Ce qui n'est évidemment pas le cas sur la figure (1.8). Plusieurs points s'écartent de la ligne de 45 degrés suggérant des changements dans le choix des individus d'un traitement à l'autre. En l'occurrence, les trois individus du bas du graphique ont tous choisi zéro choix-sûrs dans l'expérience LMF pour ensuite choisir respectivement trois, quatre et cinq choix-sûrs dans le traitement LME.

Ceci suggère alors que la distribution agrégée des choix-sûrs reste la même, mais les individus changent de comportement à l'intérieur du groupe avec la taille des mises. Pour analyser cette question plus en

[26] Chlaß and Krüger (2007) montrent que le test de rang signé de Wilcoxon, qui est un test non paramétrique peut échouer à détecter la différence entre des échantillons appariés dans le cas des petits échantillons quand certaines hypothèses sont violées notamment la continuité des variables. Ils montrent que lorsque le domaine de réponse est restreint et non continu et/ou la variable réponse est discrète, ceci peut induire une perte de puissance du test de Wilcoxon. Dans le cas de cette expérience, étant donné que le nombre de choix-sûrs est discret et contraint entre 0 et 10, ceci peut justifier le fait que les agents changent leur décision sans que le test de Wilcoxon ne puisse le détecter.

Figure 1.8: Distribution des choix-surs dans l'espace des deux traitements: LMF Versus LME



détail, considérons la régression linéaire du nombre de choix-sûrs dans le traitement LME (CS_{LME}) sur le nombre de choix-sûrs dans le traitement LMF (CS_{LMF}).

$$CS_{LME} = \beta_0 + \beta_1 CS_{LMF} + \epsilon \tag{1.36}$$

Si le comportement de prise de risque était identique dans les deux traitements, les participants auraient le même nombre de choix-sûrs dans les deux traitements. Ce qui impliquerait que $\beta_1 = 1$ et $\beta_0 = 0$. Les résultats de cette régression sont présentés dans le Tableau (1.7). La valeur estimée de β_1 est de 0,38 avec un écart-type de 0,1. Nous rejetons l'hypothèse nulle $\beta_1 = 1$ à tous les niveaux de significativité (la p-value du test de Fisher est égale zéro). Nous estimons la valeur de β_0 à 3,09 avec un écart-type de 0,48. L'hypothèse nulle $\beta_0 = 0$ est rejetée à tout seuil de significativité standard (la p-value du test de Fisher est égale zéro). Le test-joint de Fisher des deux restrictions rejette également l'hypothèse nulle $H_0 : \beta_0 = 0$ et $\beta_1 = 1$.

On conclut alors que le comportement des agents change en présence de loteries à mises élevées. Les deux mesures de préférence pour le risque, obtenues avec LME et LMF, sont donc différentes. Dans la suite, nous étudions les déterminants et la cohérence empirique de chacune de ces mesures.

Table 1.7: Estimation linéaire et test-joint de Fisher

	β_1	β_0	Test-joint de Fisher $H_0 : \beta_0 = 0$ et $\beta_1 = 1$
Estimation et test	.38***	3.09***	F(2,60) = 21.42
	(0.1)	(0.48)	Prob > F = 0.0000

*** significatif au seuil de 1%

1.9 L'analyse de la validité externe des mesures expérimentales de préférence pour le risque

La validité externe fait référence à la possibilité de transposer ou généraliser les résultats de l'expérience de terrain aux décisions économiques dans un contexte extérieur à cette expérience (Roe and Just, 2009). Dans ce papier, il s'agit de tester la capacité des préférences pour le risque (telles que mesurées dans les deux expériences des loteries LME et LMF) à prédire le choix des contrats des travailleurs.

1.9.1 La stratégie d'estimation

Afin d'évaluer la validité externe, nous testons la capacité des paramètres de préférence pour le risque (identifiés par le nombre de choix-sûrs dans l'expérience des loteries) et l'habileté du travailleur de prédire le choix des contrats conformément aux attentes théoriques.

Les préférences pour le risque et l'équivalent certain

Dans l'expérience de choix de contrat, nous avons identifié les bornes de l'équivalent certain du contrat à la pièce pour chaque travailleur. Cet équivalent certain est le salaire fixe pour lequel le travailleur est indifférent entre le contrat à la pièce et ce contrat à salaire fixe. Si l'équivalent certain est élevé, ceci implique que le travailleur est moins disposé à choisir un contrat à salaire fixe, car il accorde une grande valeur au paiement à la pièce. Si l'équivalent certain est faible, cela implique que le travailleur est prêt à accepter des faibles salaires fixes plutôt que d'être payé à la pièce. Notre technique empirique est donc d'estimer l'effet du nombre de choix-sûrs et de l'habileté sur l'équivalent certain de chaque travailleur. Si l'agent est plus averse au risque, le travailleur valorisera moins le contrat à la pièce. Ceci parce que le contrat à la pièce induit un coût du risque croissant avec l'aversion au risque (Bellemare and Shearer, 2013). Par ailleurs, un agent plus productif valorisera davantage un paiement à la performance. Soit W_i^* est l'équivalent certain du contrat à la pièce pour l'agent i, CS_i le nombre de choix-sûrs effectués dans l'expérience des loteries, K_i l'habileté du planteur et X_i le vecteur des caractéristiques individuelles du planteur. Le modèle linéaire s'écrit:

$$W_i^* = \beta_0 + \beta_1 CS_i + \beta_2 K_i + X_i' \lambda + \epsilon_i \tag{1.37}$$

Conformément aux prédictions théoriques, l'effet attendu du nombre de choix-sûrs sur le nombre de contrat à la pièce est négatif, car un agent averse au risque a une incitation plus forte à choisir un contrat à salaire fixe. À l'inverse, les agents plus productifs devraient choisir plus de contrat à la pièce. Un nombre de choix-sûrs élevé signale un individu plus averse au risque. Conformément à la prédiction 1, ceci devrait alors induire un équivalent certain plus faible. Les travailleurs plus averses au risque exigeront une prime de risque plus grande pour travailleur sous un contrat à la pièce. L'équivalent certain diminue, car il faudrait un salaire fixe plus faible pour leur assurer la même utilité que le contrat à la pièce. Ceci réduit le nombre de fois où le travailleur choisira le contrat à la pièce. Le signe espéré de β_1 est donc négatif. Par ailleurs, conformément à la prédiction 2, un travailleur plus productif devrait accorder une plus grande importance au contrat à la pièce. Le signe attendu de β_2 est positif. En effet, plus un travailleur est productif, plus son gain espéré dans le contrat à la pièce augmente. Il exigera alors une prime de risque plus faible et il faudrait un salaire fixe plus élevé pour lui procurer la même utilité que le contrat à la pièce. Son équivalent certain augmente alors et le nombre de fois où il choisit le contrat à la pièce augmente.

Les préférences pour le risque et le choix des contrats à la pièce

Pour compléter la régression linéaire sur les déterminants de l'équivalent certain de chaque travailleur, nous analysons l'impact des caractéristiques individuelles sur le nombre de contrats à la pièce choisis lors de l'expérience de choix de contrats. Le modèle de Poisson est adapté à l'estimation de ces modèles à choix discrets.
Soit N le nombre de participants, s_i le nombre de fois où le participant choisit le contrat à la pièce et Z_i le vecteur des variables explicatives. Z_i inclut le nombre de choix-sûrs dans les deux traitements, les gains précédents du travailleur servant de proxy pour l'habileté, l'âge, le genre, l'expérience, la tenure et le niveau d'éducation. La moyenne conditionnelle du modèle de Poisson s'écrit:

$$\lambda_i = exp(Z_i'\beta) \tag{1.38}$$

où β représente le vecteur des paramètres . La fonction de vraisemblance s'écrit:

$$Pr(S = s_i|Z_i) = \frac{e^{-\lambda_i(Z_i)}\lambda_i(Z_i)^{s_i}}{S_i!} \tag{1.39}$$

Les estimateurs de β sont estimés en maximisant la fonction log-vraisemblance:

$$LogL = \frac{1}{N}\sum_{i=1}^{N} log[Pr(S = s_i|Z_i)] \tag{1.40}$$

1.9.2 Les résultats de l'estimation

L'équivalent certain

Le tableau (1.8) présente les résultats de l'estimation de l'équation (1.37) par les moindres carrés ordinaires. Notre dispositif expérimental assure l'exogénéité des mesures de préférences pour le risque (Shearer, 2004).
Les deux premières colonnes du Tableau (1.8) présentent les résultats de la prédiction de l'équivalent certain en utilisant le nombre de choix-sûrs (CS) provenant de l'expérience avec des Loteries à Mises Élevées (LME) comme variable explicative. Le résultats montre que l'effet du nombre de choix-sûrs qui est le proxy de l'aversion au risque sur l'équivalent certain n'est pas statistiquement négatif et significatif pour l'expérience LME (Colonne 1). En contrôlant pour les autres caractéristique du travailleur telles que l'habileté, l'âge, le genre, l'expérience, la tenure et l'éducation, l'effet du nombre de choix sûrs reste non significatif (colonne 2). Ceci suggère donc que l'expérience LMF n'a pas permis de capter efficacement les préférences des planteurs. A l'inverse, l'effet de l'habileté sur l'équivalent certain est positif et significatif suggérant que l'habileté du travailleur augmente son équivalent certain. Ce résultat est conforme à la prédiction 2. Par ailleurs, le genre n'affecte pas la valeur de l'équivalent certain. Ce qui peut paraître contradictoire étant donné la difficulté physique de la tâche de planteur d'arbres et on devrait s'attendre à une plus grande productivité des hommes. Toutefois, la présence d'effet de sélection effectué parmi les femmes qui postulent à ce type d'emploi peut justifier ce résultat. En effet, ce sont les femmes les plus physiquement habiles et autant performantes que les hommes qui s'orientent vers ce type d'emplois. L'éducation du travailleur accroît la valeur de l'équivalent certain. Ceci s'explique par la théorie du capital humain (Becker, 1964) qui démontre que l'éducation accroît l'efficacité. Par ce canal, l'attractivité du contrat à la pièce se trouve accrue avec le niveau d'éducation. Par ailleurs, l'éducation renforce la capacité de l'agent à évaluer le risque.

Dans les deux dernière colonnes du tableau (1.8), le nombre de choix-sûrs provenant de l'expérience LME est utilisé pour prédire l'équivalent certain. On observe que l'aversion au risque estimé dans l'expérience LME a un effet négatif et significatif sur l'équivalent certain. Ce résultat est cohérent avec la prédiction 1. Il traduit l'idée que les travailleurs les plus averses au risque exigent une prime de risque plus grande pour accepter un contrat à la pièce. Dans ce cas, un contrat à salaire fixe plus faible leur garantit la même utilité que le contrat à la pièce. Ce qui se traduit par un équivalent certain plus faible.

Table 1.8: Prédire l'équivalent certain

	Variable dépendante: Équivalent certain.			
VARIABLES	LMF		LME	
Choix-sûrs	-0.551	4.635	-26.91***	-12.45*
	(7.064)	(2.943)	(7.871)	(6.347)
Habilité/100		69.06***		65.21***
		(14.29)		(13.85)
Age		0.688		1.731
		(2.909)		(3.238)
Femme		-7.049		-3.325
		(15.37)		(15.62)
Expérience		2.782		2.407
		(2.830)		(3.091)
Tenure		-1.204		-4.692*
		(2.432)		(2.365)
Education				
Post-secondarire		37.56**		65.90***
		(16.84)		(17.40)
Début Université		50.96***		34.05**
		(16.45)		(15.48)
Diplôme universitaire		50.59**		29.70
		(22.33)		(17.67)
Constante	444.9***	47.78	570.9***	145.0
	(38.20)	(73.72)	(43.51)	(103.4)
Observations	49	47	49	47

Robust standard errors in parentheses
*** p<0,01, ** p<0,05, * p<0,1

Les contrats à la pièce choisis

Le Tableau (1.9) présente les résultats de l'estimation de l'effet du nombre de choix-sûrs des expériences de loteries et l'habileté sur le nombre de fois où le participant a choisi le contrat à la pièce.
Les deux premières colonnes utilisent le nombre de choix-sûrs obtenus dans l'expérience LMF comme variables pour prédire le nombre de contrat à la pièce choisi par le travailleurs. Les résultats montrent

que, contrairement aux attentes théoriques, le nombre de choix-sûrs provenant de l'expérience LMF n'a pas d'impact significatif sur la volonté du planteur à choisir le contrat à la pièce.

A l'inverse, en utilisant le nombre de choix-sûrs provenant de l'expérience LME pour prédire le nombre de contrat à la pièce (colonne 3 et 4) , les résultats montrent que le nombre de choix-sûrs réduit le nombre de contrats à la pièce choisi. Ceci est conforme à l'attente théorique que l'aversion au risque affecte négativement la probabilité de choisir le contrat à la pièce. De plus, l'habileté du travailleur augmente la probabilité de choisir le contrat à la pièce. Ce résultat est cohérent avec celui de l'estimation linéaire et les prédictions théoriques.

Table 1.9: Predire le nombre de contrats à la pièce

	Variable dépendante: Nombre de contrats à la pièce			
VARIABLES	LMF		LME	
Choix-sûrs	-0.00150	0.0128	-0.0716***	-0.0299*
	(0.0190)	(0.00779)	(0.0179)	(0.0153)
Habileté/100		0.179***		0.166***
		(0.0336)		(0.0336)
Age		0.00184		0.00537
		(0.00706)		(0.00774)
1 si Femme		-0.0154		-0.00482
		(0.0396)		(0.0401)
Expérience		0.00800		0.00679
		(0.00735)		(0.00799)
Tenure		-0.00664		-0.0151***
		(0.00599)		(0.00544)
Education				
Post-secondaire		0.0682		0.139***
		(0.0416)		(0.0450)
Début université		0.115***		0.0751**
		(0.0407)		(0.0378)
Diplôme universitaire		0.123**		0.0696
		(0.0571)		(0.0443)
Constante	2.001***	0.965***	2.331***	1.196***
	(0.103)	(0.175)	(0.0972)	(0.249)
Observations	49	47	49	47

Robust standard errors in parentheses
*** p<0,01, ** p<0,05, * p<0,1

1.10 Conclusion

Ce papier s'intéresse à la validité interne et externe des mesures expérimentales des préférences pour le risque et leur potentiel à prédire le choix des contrats des travailleurs. Nous avons estimé les paramètres de préférences pour le risque en utilisant une séquence d'expériences de terrain dans une entreprise de taille moyenne de plantation d'arbres en Colombie Britannique. Avec les travailleurs de cette firme, nous avons réalisé une première expérience du style de Holt and Laury (2002) pour estimer les paramètres de préférence pour le risque en utilisant successivement des loteries à mises faibles (LMF), puis des loteries à mises élevées(LME). Dans le premier volet de l'expérience de mesure des préférences pour le risque avec les mises faibles, les participants peuvent gagner entre 2$ et 77$. Dans le second volet avec les mises élevées, les gains possibles sont entre 4$ et 154$. Pour compléter l'expérience, chaque travailleur doit effectuer 10 décisions dont chacune consiste à choisir entre une loterie plus risquée et une loterie moins risquée. L'une de ces décisions a ensuite été tirée aléatoirement et la loterie choisie à cette décision a été jouée pour déterminer le gain du participant. Par la suite, nous avons réalisé une expérience de choix de contrat pour étudier la validité externe des mesures de préférences pour le risque provenant de la première série d'expériences. Cette dernière expérience a consisté pour chaque travailleur à effectuer 12 décisions consistant chacune à choisir entre son contrat à la pièce habituel et un contrat à salaire fixe. Il est important de rappeler que ces choix conduisent à des conséquences réelles. En effet, après avoir complété l'expérience de choix de contrat, l'une des douze décisions effectuées par l'agent est tirée au hasard et le contrat choisi par le travailleur à cette décision lui est implémenté pendant deux jours de travail. Dans ce papier, nous développons par la suite un modèle théorique de choix de contrat. En se basant sur ce modèle, nous dérivons les deux prédictions théoriques suivantes: (i) l'aversion au risque réduit la probabilité de choisir un contrat à la pièce (ii) l'habileté accroît la probabilité de choisir le contrat à la pièce. Nous combinons les données expérimentales et les données de plantation de la firme pour tester ces prédictions.

Les résultats montrent que la distribution agrégée des préférences pour le risque est stable lorsque la taille des loteries est doublée. À l'inverse, les travailleurs ont changé leur comportement de prise de risque. Nous en déduisons donc que les mesures expérimentales de préférence pour le risque sont sensibles à l'échelle des loteries. Les mesures obtenues avec les loteries à mises faibles sont différentes des mesures obtenues avec des mises élevées. Nous avons alors analysé le potentiel de chacune d'elles à prédire le choix des contrats afin d'identifier la meilleure mesure. Les résultats montrent que les préférences pour le risque mesurées en présence d'un jeu de loterie avec des gains élevés prédisent bien le choix de contrat. En utilisant cette mesure, il apparaît que les agents qui sont les plus averses au risque ont une plus faible tendance à choisir les contrats à la pièce conformément aux prédictions théoriques. De plus, les agents les plus productifs choisissent davantage le contrat à la pièce que les agents les moins productifs. Par contre, les mesures obtenues en utilisant les loteries à mises faibles ne prédisent pas le choix des contrats. L'une des raisons serait que la hausse des montants en jeu incite les travailleurs à jouer plus sérieusement la loterie et à révéler mieux leurs préférences pour le risque. Dans ce cas, les mesures expérimentales obtenues avec des loteries à mises élevées sont plus représentatives des préférences des travailleurs.

Nos résultats impliquent que la distribution agrégée n'est pas une mesure efficace de l'effet d'échelle. Pourtant, certaines études ont analysé l'effet d'échelle en se fondant sur la distribution agrégée des préférences pour le risque lorsque la taille des mises s'accroît (Holt and Laury, 2002; Bellemare and Shearer, 2010). Ceci peut conduire à une mauvaise appréhension du comportement des individus face à une modification de la valeur monétaire des gains potentiels des loteries. De plus, au-delà du débat sur la robustesse sur les préférences mesurées dans les expériences avec gains réels comparativement aux expériences avec des gains hypothétiques (Xu et al., 2016), nos résultats suggèrent que même si les gains de la loterie sont réels, ils peuvent conduire à des mesures de préférences erronées si les montants sont faibles. La condition majeure pour que les expériences des loteries mesurent bien les préférences pour le risque est donc que les incitations financières doivent être suffisamment élevées.

Toutefois, cette étude ne permet pas d'identifier la valeur minimale des mises garantissant la validité externe des mesures expérimentales des préférences pour le risque. Par ailleurs, dans la perspective de recherches futures, il serait intéressant d'étudier en détail l'effet de la taille des mises sur le comportement de prise de risque en fonction des types d'agents ainsi que les non-linéarités potentielles de l'effet d'échelle.

Bibliography

Ackerberg, D. A. and M. Botticini (2002): "Endogenous matching and the empirical determinants of contract form," *Journal of Political Economy*, 110, 564–591.

Allen, D. W. and D. Lueck (1999): "The role of risk in contract choice," *Journal of Law, Economics, and Organization*, 15, 704–736.

Becker, G. (1964): "Human Capital National Bureau of Economic Research," *New York*.

Bellemare, C. and B. Shearer (2010): "Sorting, incentives and risk preferences: Evidence from a field experiment," *Economics Letters*, 108, 345–348.

——— (2013): "Multidimensional heterogeneity and the economic importance of risk and matching: evidence from contractual data and field experiments," *The RAND Journal of Economics*, 44, 361–389.

Belzil, C. and M. Sidibé (2016): "Internal and External Validity of Experimental Risk and Time Preferences," *Working paper*, IZA, 10348.

Cadsby, C. B., F. Song, and F. Tapon (2007): "Sorting and incentive effects of pay for performance: An experimental investigation," *Academy of management journal*, 50, 387–405.

Charness, G., U. Gneezy, and A. Imas (2013): "Experimental methods: Eliciting risk preferences," *Journal of Economic Behavior & Organization*, 87, 43–51.

Charness, G. and A. Viceisza (2011): "Comprehension and risk elicitation in the field: Evidence from rural Senegal. IFPRI discussion papers 1135," *International Food Policy Research Institute (IFPRI)*.

Chlass, N. and J. J. Krüger (2007): "Small sample properties of the Wilcoxon signed rank test with discontinuous and dependent observations," .

Dohmen, T. and A. Falk (2011): "Performance pay and multidimensional sorting: Productivity, preferences, and gender," *American Economic Review*, 101, 556–90.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*, 9, 522–550.

Eckel, C. C. and P. J. Grossman (2008): "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices," *Journal of Economic Behavior & Organization*, 68, 1–17.

Holt, C. A. and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *The American Economic Review*, 92, 1644–1655.

Jacobson, S. and R. Petrie (2009): "Learning from mistakes: What do inconsistent choices over risk tell us?" *Journal of risk and uncertainty*, 38, 143–158.

Paarsch, H. J. and B. Shearer (2000): "Piece rates, fixed wages, and incentive effects: Statistical evidence from payroll records," *International Economic Review*, 41, 59–92.

Prendergast, C. (1999): "The provision of incentives in firms," *Journal of economic literature*, 37, 7–63.

——— (2000): "What trade-off of risk and incentives?" *The American Economic Review*, 90, 421–425.

Roe, B. E. and D. R. Just (2009): "Internal and external validity in economics research: Tradeoffs between experiments, field experiments, natural experiments, and field data," *American Journal of Agricultural Economics*, 91, 1266–1271.

Shearer, B. (2004): "Piece rates, fixed wages and incentives: Evidence from a field experiment," *The Review of Economic Studies*, 71, 513–534.

Xu, S., Y. Pan, Y. Wang, A. M. Spaeth, Z. Qu, and H. Rao (2016): "Real and hypothetical monetary rewards modulate risk taking in the brain," *Scientific Reports*, 6.

Annexe

Table 1.10: Expérience LMF

Décision	Loterie A	Loterie B	E[B]-E[A]*
1	40$(1/10), 32$(9/10)	77$(1/10),2$(9/10)	-23,40$
2	40$(2/10),32$(8/10)	77$(2/10),2$(8/10)	-16,60$
3	40$(3/10), 32$(7/10)	77$(3/10),2$(7/10)	-10,00$
4	40$(4/10), 32$(6/10)	77$(4/10),2$(6/10)	-3,20$
5	40$(5/10), 32$(5/10)	77$(5/10),2$(5/10)	3,60$
6	40$(6/10), 32$(4/10)	77$(6/10),2$(4/10)	10,20$
7	40$(3/10), 32$(3/10)	77$(7/10),2$(3/10)	17,00$
8	40$(8/10), 32$(2/10)	77$(8/10),2$(2/10)	23,60$$
9	40$(9/10), 32$(1/10)	77$(9/10),2$(1/10)	30,40$
10	40$(10/10)	77$(10/10)	37,00$

*Différentiel de l'espérance de gain de la loterie B par rapport à la loterie A

Table 1.11: Expérience LME

Decision	Loterie A	Loterie B	E[B]-E[A]**
1	80$(1/10), 64$(9/10)	154$(1/10),4$(9/10)	-46,80$
2	80$(2/10), 64$(8/10)	154$(2/10),4$(8/10)	-33,20$
3	80$(3/10), 64$(7/10)	154$(3/10),4$(7/10)	-20,00$
4	80$(4/10), 64$(6/10)	154$(4/10),4$(6/10)	-6,40$
5	80$(5/10), 64$(5/10)	154$(5/10),4$(5/10)	7,20$
6	80$(6/10), 64$(4/10)	154$(6/10),4$(4/10)	20,40$
7	80$(3/10), 64$(3/10)	154$(7/10),4$(3/10)	34,00$
8	80$(8/10), 64$(2/10)	154$(8/10),4$(2/10)	47,20$
9	80$(9/10), 64$(1/10)	154$(9/10),4$(1/10)	60,80$
10	80$(10/10)	154$(10/10)	74$

**Différentiel de l'espérance de gain de la loterie B par rapport à la loterie A

Table 1.12: Expérience de choix de contrat

Decision	Option A	Je choisis A	Option B	Je choisis B
1	Contrat à la pièce		Salaire fixe de 100$ par jour	
2	Contrat à la pièce		Salaire fixe de 150$ par jour	
3	Contrat à la pièce		Salaire fixe de 200$ par jour	
4	Contrat à la pièce		Salaire fixe de 250$ par jour	
5	Contrat à la pièce		Salaire fixe de 300$ par jour	
6	Contrat à la pièce		Salaire fixe de 350$ par jour	
7	Contrat à la pièce		Salaire fixe de 400$ par jour	
8	Contrat à la pièce		Salaire fixe de 450$ par jour	
9	Contrat à la pièce		Salaire fixe de 500$ par jour	
10	Contrat à la pièce		Salaire fixe de 550$ par jour	
11	Contrat à la pièce		Salaire fixe de 600$ par jour	
12	Contrat à la pièce		Salaire fixe de 650$ par jour	

Chapter 2

Perception Bias and Risk-Incentives Trade-off : Evidence from Field Experiments

2.1 Résumé

Une explication possible à l'absence d'arbitrage risque-incitations dans les données contractuelles est que les travailleurs ont une perception incorrecte du risque auquel ils font face dans les contrats. En effet, si les travailleurs sous-estiment le risque, ils seront prêts à travailler dans un environnement très risqué en exigeant une prime de risque plus faible. Ceci réduirait alors le coût de l'incitation pour l'entreprise et supprimerait alors l'arbitrage risque-incitations qui devrait exister dans les contrats. Dans ce papier, nous analysons cette problématique en utilisant une série d'expériences de terrain auprès des travailleurs d'une firme de plantation d'arbres en Colombie-Britannique. Au cours de l'expérience, les travailleurs, payés à la pièce, doivent prédire leurs gains espérés sur une durée de 10 jours. A l'aide d'un questionnaire, nous avons construit la distribution perçue des gains quotidiens pour chaque travailleur. Ensuite, nous construisons la répartition réelle des gains de ces travailleurs en utilisant les données de paie fournies par la firme à l'issue de la saison de plantation. Notre stratégie empirique consiste à apparier la répartition de gains prédite par chaque travailleur avec la distribution réelle des gains pour évaluer l'écart potentiel entre ces deux distributions. Les résultats confirment la présence d'un biais de perception des travailleurs sur leurs gains potentiels sous le contrat de travail à la pièce. En effet, les travailleurs surestiment leurs gains journaliers moyens et sous-estiment la variabilité de leurs gains journaliers et donc le risque auquel ils font face. Par la suite, nos résultats montrent que la sous-estimation du risque accroît la probabilité des travailleurs à choisir un contrat à la pièce au détriment d'un contrat à salaire fixe.

Mots-clés: Biais de perception, risque, incitations, arbitrage , expérience de terrain

Codes JEL : C93, D81, D84

2.2 Abstract

A possible explanation for the lack of risk-incentives trade-off in observed contractual data is that workers have biased perceptions of the income risk that they face in contractual settings. For example, if workers underestimate the risk that is present, they will be willing to work in settings that are very risky for a much reduced earnings premium. This would reduce a firm's cost of implementing incentives in risky settings and would suppress the risk-incentives trade-off in observed contractual data. We investigate this issue using daily payroll data on the earnings of workers who are paid piece rates. We construct the actual distribution of earnings for individual workers in a tree-planting firm. We then elicit each worker's perceived earnings distribution, using a questionnaire. We compare the perceived distribution to the actual distribution. Our results suggest that workers overestimate their average daily earnings and underestimate the variance of daily earnings and hence the earnings risk that they face. Furthermore, we find that this under-estimation of the risk, increases workers' likelihood of choosing piece rate contracts over the fixed wage contract.

Keywords: perception bias, earning risk, risk-incentives trade-off, field experiment

JEL codes: C93, D81, D84

2.3 Introduction

The principal-agent model is based on the risk-incentives trade-off. Piece-rate contracts provide incentives by attaching workers pay to their productivity, exposing them to income risk. Risk-averse workers must be compensated for working in risky environments which increases the cost of assigning such workers to piece rate contracts. The model therefore predicts that piece rate contracts will be avoided in risky settings (Hölmstrom, 1979). Yet, empirical work often fails to uncover any relationship between risk and incentives in labour market settings (Allen and Lueck, 1992; Prendergast, 2000). One possible explanation concerns difficulties in observing risk preferences. The implied trade-off is conditional on a worker's risk preferences: risk averse workers require a premium to accept risky contracts. Yet risk neutral or risk loving workers require no premium. Many researchers have focused on risk preferences as the explanation of the missing risk-incentives trade-off but find mixed results (Allen and Lueck, 1999; Ackerberg and Botticini, 2002). Another possible explanation for the lack of risk-incentives trade-off in observed contractual data is that workers have biased perceptions of the income risk that they face. For example, if workers underestimate the risk that is present, they will be willing to work in risky settings for a much reduced earnings premium. This would reduce the firm's cost of implementing incentives in risky settings and would suppress the risk-incentives trade-off observed in contractual data.

In this paper, we address this issue by considering worker perceptions of their earnings risk within a real firm. To do so, we exploit data from field experiments conducted in a tree planting firm. Our experiments took place within a medium size tree-planting firm, located in British Columbia, Canada. Workers in this firm are hired to plant trees on recently logged parcels of land and are typically paid piece rates – daily earnings are the product of the piece rate and the number of trees planted. Since the number of trees a worker can plant depends on elements that are beyond his/her control (such as weather, the slope and the hardness of the ground), workers are exposed to daily income risk under a piece-rate contract. We investigate worker perceptions of this risk and its relation to actual risk. We elicit each worker's perceived earnings distribution, using two questions. The first question is a single value forecasting (hereafter Q1), in which workers were asked to predict their average earnings. The second question is a probabilistic forecasting question (hereafter Q2) in which workers were asked to predict the probability that their earnings would lie within a number of fixed intervals. Then, we conducted a contract choices experiment, in which workers were asked to make 12 consecutive decisions between a risky contract which pays piece rates and a series of increasing fixed wage contracts. The number of piece rate contracts a participant choose before switching from preferring piece rates to fixed wage contract depends on his/her ability, his/her perception of risk and his/her risk preferences.

To control for risk preferences, we also conducted a risk-preference-revealing experiment using the approach of Holt and Laury (2002) to measure workers' risk preferences. In this experiment, each participant had to make 10 decisions. Each decision consisted of choosing between a safe lottery and a risky lottery. The safe lottery denoted A pays either a high payoff of \$80 or a low payoff of \$64. The risky lottery denoted B pays either a high payoff of \$154 and a low payoff of \$4. From this Holt and Laury (2002) experiment, we estimated workers' risk preferences.

Our empirical strategy consists in using the worker's responses to the questionnaire on earnings perception to construct their perceived earnings distribution. Subsequently, we construct the actual distribution of earnings for each worker based on the firm's payroll records and observed daily earnings.

We compare the perceived distribution to the actual distribution in order to evaluate the presence of perception biases over earnings risk and its effects on the risk-incentive trade-off in incentive contracts. Our results suggest that workers overestimate their average daily earnings by between 7,89 percent and 11,3 percent. In contrast, they underestimate the standard deviation of their daily earnings by around $150 and hence the earnings risk that they face. We explore the effect of this observed perception bias on the choice between a piece rate and a fixed wage contract.

With regards to the ability of risk perception bias to predict contract choices, our results suggest that underestimating risk increases a worker's likelihood of choosing piece rate contracts over fixed wage contracts. These results demonstrate the usefulness of perception bias as an explanation for the missing risk-incentives trade-off in contractual data.

Issues related to biased perception of risks have been addressed in other settings, notably the risk due to smoking (Viscusi, 1990, 2016), the self-perception of weight and its association with weight-related behaviors (Rahman and Berenson, 2010), the perceived personal immunity to the consequences of drinking alcohol and substance abuse (Hansen et al., 1991), the perception of mortality risk due to different causes such as road accidents, suicide, cancer, heart disease, stroke (Lichtenstein et al., 1978; Gegax et al., 1991; Viscusi et al., 1997), the perception of job-related risks (Gegax et al., 1991) and the perception of uncertain threats posed by several drinking water risks (Viscusi et al., 2019). Viscusi (1993) and Viscusi et al. (1979) focused on health and safety risk (illness, injury, death...) in job markets to estimate a compensating wage differential for risky jobs. Using a state-dependent utility model, Viscusi (1993) and Viscusi et al. (1979) show that, regardless to their risk preferences, workers demand compensating differentials for risk since one would rather be healthy than injured or dead. The key assumption leading to this result is that the utility of being healthy is higher than the utility of being injured. In this context, Viscusi (1993) found that the compensating risk premium is increasing with the risk level for all the workers. So, if a worker underestimates risk, he/she will be more likely to choose a risky job[1].

In contrast, our theoretical model uses a state independent utility close to Shearer (2004) and Bellemare and Shearer (2013). We show that, in the context of financial risk, the risk premium increases with the risk only if the worker is risk averse. Hence, underestimating earnings risk increases the utility of the risky contract only if workers are risk averse. This implies that risk averse workers will accept risky incentive contracts with a lower earnings premium than they would if they perceived earnings risk correctly. The intuition behind this result is that while health and safety risks always reduce workers' utility, the effect of earnings risk on utility is different. Indeed, a production shock reduces earnings (and the utility) when the production shock is negative and increases earnings (the utility) when the production shock is positive. Being subject to such a risk, the compensating differentials for the perceived risk depends on a worker's attitudes toward risk. *Ceteris paribus*, the risk averse worker is attracted by a risk-free environment, and will require incentives to work in a risky environment while a risk loving worker does not. In order to test for the role of risk preferences, we interact the measured risk preferences with the perception bias in the empirical estimation to predict contract choices. Our evidence on the interaction effect between risk perception and risk preferences is less conclusive and requires further study. More data will be needed to test for these interaction effects precisely.

[1] When we apply a state independent utility in the context of Viscusi (1993), risk preferences matter in the wage-risk trade-off (see Appendix 4)

The remainder of this paper is structured as follows. Section 2 presents the context of the experiment. Section 3 develops the theoretical model from which we derive the effect of perception bias on the equilibrium risk premium. Section 4 presents the data and the identification strategy. Section 5 discusses the results of our estimation. Finally, section 6 provides a conclusion.

2.4 Experimental Context

2.4.1 The firm

Our experiments took place within a tree-planting firm, located in British Columbia, Canada, during the planting season 2016. Workers in this firm are hired to plant trees on recently logged parcels of land and are typically paid piece rates – daily earnings are the product of the piece rate and the number of trees planted. Since the number of trees a worker can plant depends on elements that are beyond his/her control (such as weather, the slope and the hardness of the ground) workers are exposed to daily income risk under a piece-rate contract. The ability of workers to perceive this risk is the basis of the risk-incentives trade-off of agency theory. Our experiment is designed to analyzed worker's perception of that risk and the implications of perception bias -if any- on workers self-selection in contracts.

2.4.2 The field experiments

We conducted field experiments to evaluate planters' ability to predict their choices risk under the piece rate contract and the effect of the perception bias on workers contract choices while controlling for their risk preferences.

The questionnaire to elicit earnings perceptions

Our questionnaire is close to that of Dominitz and Manski (1997) who analyzed the ability of students to forecast their expected returns to schooling. In the first question, workers were asked to predict their average earning on a typical day. In the second question, workers were asked to indicate over a period of 10 working days, the number of days he/she expects to receive an income included in different intervals of \$50[2] (See Table 2.8). Dominitz and Manski (1997) show that the internal consistency of the respondents' forecasting can be accessed by checking the existence of three possible errors:

1. The probability reported by a participant does not fall within the unit interval.
2. The sequence of probabilities reported by a participant violates monotonicity of the cumulative distribution function.
3. A participant reported a distribution of probability that is logically inconsistent with the previously reported average earnings in Q1.

The results from our questionnaire suggest that no participant committed an error of the first or second type. For the third type of error, Figure (2.1) gives a spatial view of the perceived average earnings from Q1 and Q2. In Figure (2.1), we observe that the observations are grouped along the 45 degree line suggesting that two measures are very close.

[2] Except the first interval which is from 0 to 100

The contract choices experiment

We conducted a contract choices experiment with the same planters. During this experiment, workers were asked to make 12 consecutive decisions between a risky contract which pays piece rates and a fixed wage contract (See Table 2.9). At the first decision, each worker chooses between earning a fixed wage of $100 per day and the regular piece rate contract paid by the firm. At this point, only a worker with very low productivity and/or extreme risk aversion would choose the fixed wage contract over the piece rate contract[3]. The piece rate contract is always the same (the regular contract paid by the firm). The fixed wage is increasing while the participant moves down the sheet and a participant will eventually switch to choose the fixed wage contract. We define the certainty equivalent of the piece rate as the earnings level at which the planter shifts from preferring the piece-rate contract to the fixed wage contract. The result of this experiment allows us to evaluate how perception bias influences the risk incentive trade-off namely the impact of perception bias on workers choice between the safe contract and the risky contract.

The risk-preference-revealing experiment

We conducted a risk-revealing experiment using the approach of Holt and Laury (2002) to measure workers' risk preferences[4]. To complete the experiment, each participant had to make 10 decisions. Each decision consisted of choosing between a safe lottery and a risky lottery. The safe lottery denoted A pays either a high payoff of $80 or a low payoff of $64. The risky lottery denoted B pays either a high payoff of $154 and a low payoff of $4. The final earnings was determined by chance. The sheet of the risk-preference-revealing experiment is presented in Table (2.10). At the first decision, the probability of the low payoff is 90 percent and the probability of the high payoff is only 10 percent. Therefore, the expected payoff from lottery A is 46.8 higher than for lottery B. At this decision, only an extreme risk loving and/or low ability worker will choose lottery B. However, when moving down the sheet, the probability of a high payoff increase, increasing the expected payoff of risky lottery. At the final decision, all the participants would choose lottery B because it yields the high payoff of 154$ with a probability equals 1. All the participants would eventually shift at a certain decision from preferring lottery A to lottery B. As in Holt and Laury (2002), we measure the participant's risk aversion by the number of time he/she chooses the safe lottery (lottery A) before switching to choose the risky lottery (lottery B).

2.5 Theoretical model

In this section, we present a parsimonious model to analyze the implications of perception bias for workers' risk premia. We build our model around the conceptual framework developed by Shearer (2004) and Bellemare and Shearer (2013) in the same context of tree-planting. However, while Shearer

[3]Given that the average choices of the piece rate is around $425

[4]Actually, we conducted two risk-preference-revealing experiments. The first experiment use low stake lotteries while the second experiments use high stake lotteries. In these two experiments, the participant make 10 decisions between a safe lottery and a risky lottery. During the low stake lotteries experiment, the safe lottery pays either a high payoff of $40 or a low payoff of $32 while the risky lottery pays either a high payoff of $77 and a low payoff of $2. During the high stake lotteries experiment, all the payoffs are doubled. Previous analysis revealed that the high stake lotteries experiment gives better measures of individuals' risk preferences. Hence, in this paper, we report the experiment with high payoffs that will be use for our analysis to control for workers' risk preferences

(2004) focuses on the effect of piece rates on workers productivity and Bellemare and Shearer (2013) measure the cost of risk and the benefits of matching heterogeneous workers to risk levels within a firm, we design our model to predict the effect of risk perception bias on contract choices.

2.5.1 Technology

Consider workers indexed by $i \in \{1...N\}$, who are randomly allocated to plant trees on different blocks indexed $j \in \{1...J\}$. Let Y_{ij} denote the daily productivity of a worker i on the block j. This productivity depends on the worker's effort E_{ij} and a productivity shock $s_j \in S$.

$$Y_{ij} = e_{ij}s_j \tag{2.1}$$

The productivity shock s_j is a disturbance factor beyond the worker's control which captures the planting conditions such as the ground hardness, the soil quality, the soil type and temperature. Let assume that $ln(s_j) \sim N(\mu_j, \sigma^2)$.

2.5.2 Preferences

Under a piece rate contract, worker $i's$ daily earnings, w_{ij}, depend on the number of trees he/she plants Y_{ij} and the piece rate r_j on the block j.

$$w_{ij} = r_j Y_{ij} \tag{2.2}$$

We assume that worker $i's$ preferences are represented by the following CRRA utility function:

$$U(w_{ij}, e_{ij}) = \begin{cases} \frac{1}{\delta_i}[w_{ij} - C(e_{ij})]^{\delta_i} & \text{if } w_{ij} > C(e_{ij}) \\ -\infty & \text{otherwise,} \end{cases} \tag{2.3}$$

where $C(e_{ij}) = \frac{1}{\eta}k_i e_{ij}^{\eta}$ represents the effort cost, δ_i is the risk preferences parameter and k_i captures worker $i's$ ability.

2.5.3 Timing

For a given block, j, to be planted:
(i) nature chooses the average shock μ_j and the variance σ^2;
(ii) the firm observes μ_j and σ^2 and propose the piece rate r_j to the worker ;
(iii) the worker observes μ_j and σ^2 ;
(iv) the worker accepts or rejects the contract r_j ;
(v) conditional on accepting the contract, the worker draws a value s_j from the distribution of S_j and chooses an effort level e^*_{ij}, producing Y_{ij};
(vi) the firm observes Y_{ij} and pays earnings w_{ij}.

2.5.4 Indirect utility

Under the piece rate contract, each worker i chooses his/her optimal effort e^*_{ij} on the block j to maximize his/her utility $U(w_{ij}, e_{ij})$. The optimal effort is then:

$$e^*_{ij} = \left[\frac{r_j s_{ij}}{k_i}\right]^{\gamma} \text{ where } \gamma = \frac{1}{\eta - 1}. \tag{2.4}$$

The optimal effort e^*_{ij} does not depend on risk preferences since s_j is observed before workers choose their effort. The second order condition: $\frac{\partial^2 U}{\partial E^2_{ij}} < 0$[5] ensures that such an optimum exists.

Given the piece rate r_j, the productivity shock s_{ij} and his/her optimal effort e^*_{ij}, worker i is paid a wage :

$$w(r_j, s_{ij}) = \frac{r_j^{\gamma+1}}{k_i^{\gamma}} s_{ij}^{\gamma+1} \tag{2.5}$$

and his/her effort cost is,

$$C(e_{ij}) = \frac{\gamma}{\gamma+1} \frac{r_j^{\gamma+1}}{k_i^{\gamma}} s_{ij}^{\gamma+1} \tag{2.6}$$

The indirect utility $V(s_{ij})$ is given by:

$$V(s_{ij}) = \frac{1}{\delta_i} \frac{r_j^{\delta_i(\gamma+1)}}{(\gamma+1)^{\delta_i} k_i^{\delta_i \gamma}} s_{ij}^{\delta_i(\gamma+1)} \tag{2.7}$$

By using the properties of the log-normal functions on Eq (2.5), Eq.(2.6) and(2.7) respectively, expected earnings is given by:

$$E\left[w_{ij}(k_i, k_h, \mu_j, \sigma)\right] = \frac{r_j^{\gamma+1}}{k_i^{\gamma}} exp^{(\gamma+1)\mu_j + \frac{1}{2}(\gamma+1)^2\sigma^2}, \tag{2.8}$$

his/her expected cost of effort is,

$$E\left[c(e_{ij})\right] = \frac{\gamma}{\gamma+1} \frac{r_j^{\gamma+1}}{k_i^{\gamma}} exp^{(\gamma+1)\mu_j + \frac{1}{2}(\gamma+1)^2\sigma^2} \tag{2.9}$$

and his/her expected indirect utility is given by:

$$E\left[V_{ij}(r_j, \mu_j, \sigma)\right] = \frac{1}{\delta_i} \frac{r_j^{\delta_i(\gamma+1)}}{(\gamma+1)^{\delta_i} k_i^{\delta_i \gamma}} exp^{\delta_i(\gamma+1)\mu_j + \frac{1}{2}(\gamma+1)^2 {\delta_i}^2 \sigma^2} \tag{2.10}$$

The firm chooses the piece rate r_j to satisfy the workers' participation constraint. Let $\bar{w}$ denotes the alternative wage on the market and let h denote the marginal worker who is indifferent between his/her alternative and the piece rate contract. The piece rate solves:

$$\frac{1}{\delta_h} \frac{r_j^{\delta_h(\gamma+1)}}{(\gamma+1)^{\delta_h} k_h^{\delta_h \gamma}} \exp^{(\delta_h(\gamma+1)\mu_j + \frac{1}{2}(\gamma+1)^2 {\delta_h}^2 \sigma^2)} = \frac{1}{\delta_h} \bar{w}^{\delta_h} \tag{2.11}$$

which implies

$$r_j^{\gamma+1} = \bar{w}(\gamma+1) k_h^{\gamma} \exp^{-\left[(\gamma+1)\mu_j + \frac{1}{2}(\gamma+1)^2 \delta_h \sigma^2\right]} \tag{2.12}$$

[5] $\frac{\partial^2 U}{\partial E^2_{ij}} = -\frac{1}{\gamma} k_i^{\gamma} (r_j s_{ij})^{1-\gamma}$

By substituting $r_j^{\gamma+1}$ from (2.12) into (2.10), the equilibrium expected utility for worker i on the block j :

$$E\left[V_{ij}^r(\delta_i,k_i)\right] = \frac{1}{\delta_i}\left[\bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma}\right]^{\delta_i} \exp^{\frac{1}{2}\delta_i(\gamma+1)^2(\delta_i-\delta_h)\sigma^2} \tag{2.13}$$

Replacing $r_j^{\gamma+1}$ from (2.12) into (2.9), gives the equilibrium effort cost for worker i on the block j under piece rates is:

$$E\left[c(e_{ij})\right] = \gamma\bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma} \exp^{\frac{1}{2}(\gamma+1)^2(1-\delta_h)\sigma^2} \tag{2.14}$$

By substituting $r_j^{\gamma+1}$ from (2.12) in (2.8), the equilibrium expected income for worker i on the block j under piece rates,

$$E\left[w_{ij}(\delta_i,k_i)\right] = \bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma} \exp^{\frac{1}{2}(\gamma+1)^2(1-\delta_h)\sigma^2} + E\left[C(e_{ij}^*)\right] \tag{2.15}$$

Worker $i's$ equilibrium expected income in Eq (2.15) has two parts : The first part represents the earnings paid to compensate the cost of risk under the contract j. This cost depends on σ^2 the level of risk , (k_h, δ_h) the marginal workers parameters and k_i the worker $'i$ ability. The second part is the earnings paid to compensate worker's effort.

2.5.5 Equilibrium risk premium

Let $\bar{W}_{ij}$ denote, the certainty equivalent income which represents the fixed wage that provides worker i the same level of utility as a piece rate contract, holding his effort cost constant at piece rate effort levels.

$$\frac{1}{\delta_i}\left[\bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma} \exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-\delta_h)\sigma^2}\right]^{\delta_i} = \frac{1}{\delta_i}\left[\bar{W}_{ij} - E\left(C(e_{ij}^*)\right)\right]^{\delta_i} \tag{2.16}$$

which implies

$$\bar{W}_{ij} = \bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma} \exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-\delta_h)\sigma^2} + E\left[C(e_{ij}^*)\right] \tag{2.17}$$

The equilibrium risk premium RP_{ij} is the amount a worker will pay in order to avoid risk, while providing the same effort level as under the piece rate contract. The equilibrium risk premium is then obtained by subtracting the certainty equivalent income $\bar{W}_{ij}$ (Eq. 2.17) from the equilibrium expected earnings (Eq. 2.15). This yields

$$RP_{ij} = \bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma}\left(\exp^{\frac{1}{2}(\gamma+1)^2(1-\delta_h)\sigma^2} - \exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-\delta_h)\sigma^2}\right) \tag{2.18}$$

Inspection of (2.18) indicates that the risk premium consists of two parts: the first part $\bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma}$ reveals that the risk premium is positively proportional to planting ability suggesting that productive workers require higher risk premia to work under a piece rate contract. The second part in parentheses reveals that the risk premium is negatively correlated with worker risk preferences. Risk averse workers require a positive risk premium to work under a risky contract ($\delta_i < 1$) while the risk premium is negative for risk loving workers ($\delta_i > 1$) (see demonstration in Appendix 2). Equation (2.18) also reveals that the risk premium is zero in the risk-free contract ($\sigma^2 = 0$).

Proposition 2.5.1 *Consider a worker i who has a biased perception of the risk factor σ^2. This worker perceived a risk level equals $\sigma^2 + \epsilon_{ij}$ where ϵ_{ij} is the perception bias on the risk. If $\epsilon_{ij} > 0$, worker i overestimates the earnings risk and if $\epsilon_{ij} < 0$, worker i underestimates the earnings risk.*

Given ϵ_{ij}, the risk premium required to attract the worker to piece rates is obtained by replacing the perceived risk in Equation (2.18):

$$RP_{ij} = \bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma}\left(\exp^{\frac{1}{2}(\gamma+1)^2(1-\delta_h)(\sigma^2+\epsilon_{ij})} - \exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-\delta_h)(\sigma^2+\epsilon_{ij})}\right) \tag{2.19}$$

In Equation (2.19), the risk premium depend on the perception bias as well. From Eq (2.19),

$$\frac{\partial RP_{ij}}{\partial \epsilon_{ij}} = \frac{1}{2}\bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma}(\gamma+1)^2\exp^{\frac{1}{2}(\gamma+1)^2(1-\delta_h)(\sigma^2+\epsilon_{ij})}\left[(1-\delta_h)-(\delta_i-\delta_h)\exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-1)(\sigma^2+\epsilon_{ij})}\right] \tag{2.20}$$

A close inspection of Eq. (2.20) gives the following predictions (see demonstration in Appendix 3):

- $\frac{\partial RP_{ij}}{\partial \epsilon_{ij}} > 0$ if $\delta_i < 1$. For risk averse workers, overestimating risk reduces the value of a piece rate contract in favour of fixed wage contracts.
- $\frac{\partial RP_{ij}}{\partial \epsilon_{ij}} < 0$ if $\delta_i > 1$. For risk loving workers, overestimating risk increases the value of piece rate contracts.
- $\frac{\partial RP_{ij}}{\partial \epsilon_{ij}} = 0$ if $\delta_i = 1$. The neutral worker is not affected by risk perception bias.

We use payroll and experimental data to, first, test for perception bias in risk levels and second, to see if this perception bias has an effect on worker contractual choice.

2.6 Data and methodology

2.6.1 Perceived and actual earnings distributions

In order to elicit the parameters of each workers' perceived distribution of their earnings (the average and the standard deviation), we rely on workers' response to Q2. In Q2, workers' were asked to predict over a typical period of 10 working days, how many days they expect their earnings will lie in a given interval indexed by j. For an individual i, let d_ij denote the number of times he/she indicated that his/her earnings will fall in interval j. We use INF_j and SUP_j to denote the lower bound and the upper bound of interval j, and n denotes the number of intervals. The perceived average earnings $Y^*_{i,a}$ is calculated as follows :

$$Y^*_{i,a} = \frac{1}{n}\sum_{j=1}^{n} d_{ij}c_j \tag{2.21}$$

where $c_j = \frac{INF_j+SUP_j}{2}$ is the center of interval j.

The perceived variance $Y^*_{i,v}$ is :

$$Y^*_{i,v} = \frac{1}{n}\sum_{j=1}^{n} d_{ij}(c_j - Y^*_{i,a})^2 \tag{2.22}$$

The perceived earnings standard deviation is the square root of the perceived variance.

We constructed the actual of earnings for each individual worker i using payroll data from the firm. In order to obtained comparable estimates, we use the following bootstrap sampling methodology to calculate the actual average and the actual standard deviation for each worker i.

1. We randomly draw a sample of 10 days of earnings by worker i from workers' payroll data.
2. We calculate and store the actual average and the actual standard deviation of this sample.
3. We repeat steps (1) and (2) 1000 times.
4. We calculate the average and the standard deviation as the mean of these 1000 replications.

2.6.2 Descriptive Statistics

Table (2.1) presents characteristics of the participants. Participant's ages range between 21 and 47 years old and 44.3 percent of them are women. With regards to education, 68.1 percent of planters in this firm have a university degree. In addition, this cohort of workers includes experienced workers with up to 28 years of experience, but also planters with only 2 years of experience. The average tenure of workers is just over 4 years.

Table 2.1: Workers characteristics

VARIABLES	N	Mean	Sd	min	max
Age	47	28.91	4.169	21	47
Female	47	0.426	0.500	0	1
Experience	47	8.511	4.510	2	28
Tenure	47	4.106	4.264	1	26
Education					
Some secondary	47	0.0426	0.204	0	1
Secondary	47	0.0213	0.146	0	1
Some university	47	0.255	0.441	0	1
University degree	47	0.681	0.471	0	1

Table (2.2) presents descriptive statistics on workers' actual and predicted earnings. The results suggest that, on average, when answering Q1, workers predicted that their daily earnings will range between \$300 and \$700 with a mean at \$458.7 and a standard deviation of \$89.07. When answering Q2, workers predicted that their daily earnings will range between \$315 and \$725 with a mean at \$473.4 and a standard deviation at \$88.23.

With respect to the actual earnings, on average, workers earned between \$280.7 and \$698.3 with an average at \$426.2, which is lower than the two predicted means. As far as the earnings risk is concerned, workers predicted that their earnings standard deviation will range between \$0 to \$75 with an average at \$40.8 and a standard deviation at \$13.9. However, the actual standard deviation of earnings appears to be in a much larger range (between \$119 and \$291) with an average roughly higher (\$191.5). During the survey, one participant indicates that his earnings does not vary across the intervals. We took out this individual to analyze the sensibility of our result. Table (2.11) reports the estimated perception bias. The perception bias on the average earnings and the standard deviation remain almost identical to Table (2.2) . Figure (2.5) present histograms of perceived and actual earnings by worker. Observational evidence from these graphs clearly indicates that the distribution of actual earnings is more flat and right-centered for most of workers compared to their actual earnings distribution.

Table 2.2: Descriptive Statistics of Perceived and Actual data

VARIABLES	N	mean	sd	min	max
Perceived data					
Perceived average of earnings Q1	47	458.7	89.07	300	700
Perceived average of earnings Q2	47	473.4	88.23	315	725
Perceived standard deviation	47	40.8	13.9	0	75
Actual data					
Number of days worked by a planter	47	42.11	4.944	26	50
Number of trees Planted	47	469.6	306.9	10	1 660
Actual average of earnings	47	426.2	92.88	280.7	698.3
Actual standard deviation of earnings	47	191.5	38.57	119	291

Table (2.3) presents the quality of the predictions based on the difference between the perceived earnings and the actual earnings. The results indicate that a large majority of the participants (78.7 percent in Q1 and 95.8 percent in Q2) overestimated their actual average earnings. Figure (2.2) and (2.3) confirm that the perceived average earnings from Q1 and Q2 respectively are higher the actual average of earnings[6]. With respect to standard deviation, all workers underestimate the standard deviation of their earnings (see Figure 2.4). Together, these results suggest that workers of this firm have incorrect perceptions of the earnings risk they face under the piece rate contract.

Table 2.3: Actual vs predicted earnings: quality of forecasting

	Under-estimation		*Over-estimation*		
VARIABLES	N	%	N	%	Total
Perceived average earnings (Q1)	10	21.3	37	78.7	47
Perceived average earnings (Q2)	2	4.2	45	95.8	47
Perceived standard deviation	47	100	0.00	0.00	47

2.6.3 Identification of Perception Bias

Consider $\hat{Y}_{i,a}$ the perceived average earnings, $\hat{Y}_{i,v}$ the perceived risk, $Y^*_{i,a}$ the actual earnings and $Y^*_{i,v}$ the actual risk. As documented in the literature (McDonald, 1973; Eddy and Seifert, 1992; Manski, 2004), we use two measures of the perception bias: the absolute perception bias and the relative perception bias.

[6] y represents the 45 degree line

Figure 2.1: Average perceived earnings Q1 vs Q2

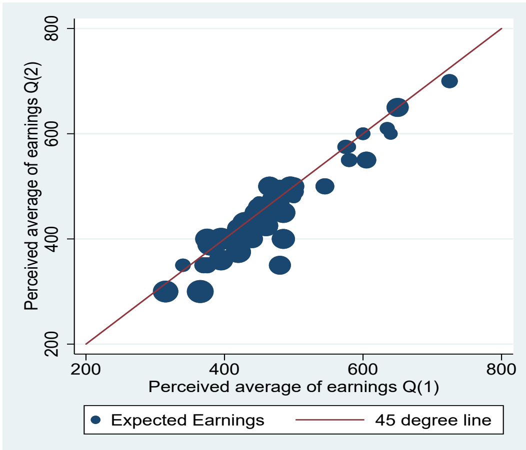


Figure 2.2: Average perceived earnings Q1 vs Average actual earnings

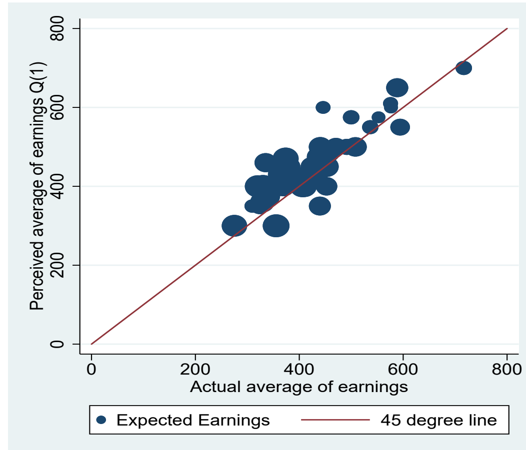


Figure 2.3: Average perceived earnings Q2 vs Average actual earnings

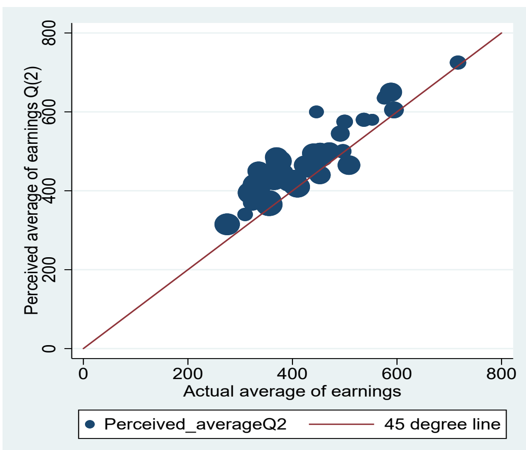
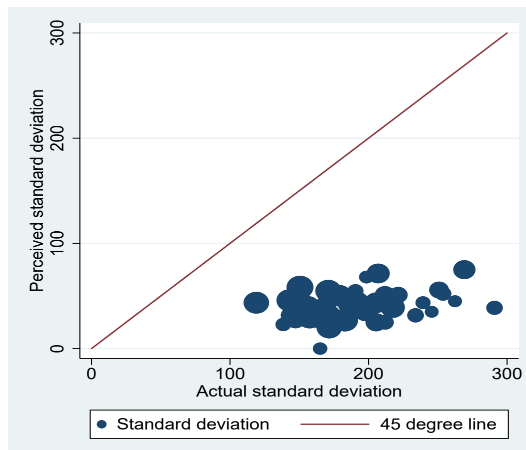


Figure 2.4: perceived vs actual standard deviation of earnings



The absolute perception bias $\tilde{y}_{i,s}$ ($s = \{a, v\}$) is the difference between the predicted earnings (the predicted standard deviation) $\hat{Y}_{i,s}$ and the actual earnings (the actual standard deviation) $Y_{i,s}$,

$$\tilde{y}_{i,s} = \hat{Y}_{i,s} - Y^*_{i,s} \tag{2.23}$$

We use the measure of perception bias for each worker to calculate the average perception bias among workers, using the equation:

$$\tilde{y}_{i,s} = \beta_s + \mu_{i,s} \tag{2.24}$$

Here, the intercept β captures the average perception bias among workers and $\mu_{i,s}$ is an error term. If $\beta_s = 0$, there is no perception bias on average. If $\beta_s > 0$ workers overestimate their average earnings (risk). If $\beta_s < 0$, workers underestimate their average earnings (risk).

The relative perception bias λ_i is written as :

$$\lambda_{i,s} = \frac{\hat{Y}_{i,s} - Y^*_{i,s}}{Y^*_{i,s}} = \frac{\tilde{y}_{i,s}}{Y^*_{i,s}} \tag{2.25}$$

We estimate the following linear model to evaluate the average relative perception bias,

$$\lambda_{i,s} = \alpha_s + \nu_{i,s} \tag{2.26}$$

The relative perception bias is interpreted in the same manner as the absolute perception bias. If $\alpha_s = 0$, there is no perception bias on average. If $\alpha_s > 0$ workers overestimate their average earnings (risk). If $\alpha_s < 0$, workers underestimate their average earnings (risk).

2.7 Results

2.7.1 Estimates of Perception Bias

Table (2.4) presents the estimates from Eq. (2.23). The results suggest that for Q1, the workers' prediction of their average earnings is $28.74 higher than the actual average earnings with a standard deviation of $6.3. In Q2, the perception bias on the average earnings is also positive ($43.42) with a standard deviation of $5.14. This result suggests that, in both experiments, workers tend to overestimate the average earnings. In contrast, we observe that workers underestimate the standard deviation of their earnings. In fact, the predicted standard deviation is, on average, $150.7 lower than the actual standard deviation.

Table 2.4: Absolute perception bias

VARIABLES	Predicted average earnings Q1	Predicted average earnings Q2	Predicted Standard deviation
Absolute Bias	28.74***	43.42***	-150.7***
	(6.299)	(5.147)	(5.327)
Observations	47	47	47

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table (2.5) presents the estimates from Eq.(2.26). The results suggest that the predicted average earnings in Q1 is 7.52 percent higher than the actual average earnings and the predicted average earnings in Q2 is 11 percent higher than the actual average earnings. We find that the average earnings, while over estimated, is of the same order of magnitude. In contrast, the predicted standard deviation is 78.3 percent lower than the actual standard deviation, which represents a huge difference. From these results, the key finding is that, while workers slightly overestimate their average daily earnings, they largely underestimate the standard deviation of their earnings and hence the risk that they face.

2.7.2 Observable heterogeneity

In this section, we analyze the link between the perception bias and workers' individual characteristics. We estimate the following equation:

$$PB_i = \alpha + X_i'\beta + \epsilon_i \tag{2.27}$$

where PB_i is the perception error, and X_i the vector of individual characteristics including age, gender, experience, tenure, and education level. We also control for worker's risk preferences measured from

Table 2.5: Relative perception bias

VARIABLES	Predicted average earnings Q1	Predicted average choices Q2	Predicted Standard deviation
Relative Bias	0.0752***	0.110***	-0.783***
	(0.0160)	(0.0137)	(0.0110)
Observations	47	47	47
R-squared	0.000	0.000	0.000

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

the risk-preference-revealing experiment mentioned above. Table (2.6) reports the estimates of Eq. (2.27). We find that the perception bias on average earnings is not systematically related to individual characteristics. This suggests that all workers overestimate their average earnings during the straight forecasting (Q1).

With respect to the perception bias on risk, women are less likely to underestimate the level of risk (see Table 2.6). This result is consistent with Savage (1993) who finds that women tend to predict a higher risk level than men. Surprisingly, we find that having a university degree increases the underestimation of the earnings risk.

Table 2.6: Perception bias and individual characteristics

	Absolute Perception bias			Relative perception bias		
	AE(Q1)	AE(Q2)	Earnings Std dev.	AE(Q1)	AE(Q2)	Earnings Std dev.[7]
Age	-2.501	-2.094	0.215	-0.00560	-0.00466	-0.00184
	(2.883)	(2.996)	(1.762)	(0.00690)	(0.00749)	(0.00550)
Risk aversion	-3.188	0.310	1.475	-0.0111	-0.000679	-0.00149
	(5.606)	(4.450)	(4.114)	(0.0152)	(0.0120)	(0.0124)
Female	4.730	-3.051	32.60***	0.0325	0.0180	0.00405
	(14.12)	(12.28)	(7.149)	(0.0370)	(0.0331)	(0.0251)
Experience	-0.743	-2.840	-3.795**	-0.00237	-0.00832	-0.00705
	(3.216)	(2.884)	(1.544)	(0.00756)	(0.00720)	(0.00519)
Tenure	0.912	3.575	-1.167	0.00111	0.00747	0.00127
	(3.468)	(3.298)	(1.701)	(0.00834)	(0.00818)	(0.00491)
Secondary degree	38.61	25.07	-42.38**	0.0826	0.0669	-0.0273
	(61.76)	(15.23)	(19.39)	(0.152)	(0.0414)	(0.0420)
Some University	31.10	-15.33	-33.45*	0.0575	-0.0480	-0.00811
	(58.82)	(10.16)	(16.62)	(0.144)	(0.0290)	(0.0312)
University degree	39.49	-0.490	-38.22**	0.0705	-0.0241	-0.0108
	(59.50)	(12.40)	(16.62)	(0.145)	(0.0332)	(0.0277)
Constant	81.16	117.0*	-105.3**	0.228	0.307*	-0.660***
	(91.65)	(66.62)	(48.84)	(0.232)	(0.170)	(0.149)
Observations	47	47	47	47	47	47
R-squared	0.091	0.125	0.711	0.115	0.189	0.193

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.7.3 Impact of perception bias on contract choices

In table (2.7), we take advantage of the data provided by our contract choices experiment conducted simultaneously in the same firm, to analyze the effect of perception bias on contract choices.

We estimate the impact of underestimating risk[8] on the number of piece rate contracts a worker has chosen over fixed wages. The results in table (2.7) indicate that the more is the underestimation of risk, the larger piece rate contracts the planter chooses. Moreover, Table (2.7) indicates that risk aversion reduces workers likelihood to choose piece rate contract while ability increases that probability. In order to test the theoretical prediction that underestimating risk increases the probability to choose a piece rate contract over a fixed wage contract only for risk averse workers, we include a risk averse dummy interacting with risk perception bias. However, we find that this interaction is not significant while risk underestimation variable alone remains significant.

Table 2.7: Perception bias and contract choices

	The dependant variable in the number of piece rates chosen					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Underestimation of risk	1.05e-05***	1.05e-05***	1.08e-05***	9.39e-06***	6.52e-06***	6.12e-06***
	(9.97e-07)	(1.08e-06)	(1.02e-06)	(1.46e-06)	(2.02e-06)	(2.08e-06)
Perception bias x Risk averse		3.55e-07	3.03e-07	-1.50e-06	-8.70e-07	-2.07e-06
		(1.21e-06)	(1.13e-06)	(1.54e-06)	(1.43e-06)	(1.58e-06)
Bias on average earnings Q1			0.000979**	0.000808	0.000825*	0.000484
			(0.000476)	(0.000494)	(0.000452)	(0.000406)
Bias on average earnings Q2			-0.000816	-0.000550	-0.000445	2.33e-05
			(0.000553)	(0.000726)	(0.000620)	(0.000634)
Risk aversion				-0.0307	-0.0285	-0.0490**
				(0.0196)	(0.0178)	(0.0234)
Ability					0.00152**	0.00115*
					(0.000643)	(0.000669)
Female						-0.00499
						(0.0403)
Tenure						-0.0160*
						(0.00873)
Experience						0.0106
						(0.00834)
Age						0.00406
						(0.00717)
University degree						0.00601
						(0.0401)
Constant	1.613***	1.618***	1.614***	1.786***	1.645***	1.648***
	(0.0453)	(0.0546)	(0.0520)	(0.123)	(0.121)	(0.210)
Observations	47	47	47	47	47	47

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.8 Conclusion

In this paper, we have measured the perception bias on the earning risk and the average earnings among workers who are paid piece rates . Further, we analyzed the effect of this perception bias on the risk-incentives trade-off in contractual data. We exploited data from field experiments that evaluate workers' ability to predict their earnings under a piece rate contract. First, we elicited each worker's perceived earnings , using two questions. In the first question, workers were asked to predict their perceived average earnings. In the second question, workers were asked to predict the probability that their earnings would lie in a given interval. Subsequently, we constructed the distribution of

[8] For a planter i, the underestimation of risk is the difference between the actual earnings risk and the perceived risk. Underestimation of risk = Actual earnings risk-Perceived earnings risk

actual earnings for individual workers based on their daily production. We compared the perceived to the actual in order to evaluate the presence of perception bias on earnings risk and its effects on workers risk-incentive trade-off. In the second experiment, we allowed workers to choose between a piece rate contract and a fixed wage contract. This enabled us to test the effect of perception biases on contract choices. In a third field experiment, we use Holt and Laury (2002)'s approach to measure workers risk preferences.

Our results suggest that workers overestimate their average daily earnings between 7,52 percent and 11 percent. In contrast, they underestimate the standard deviation of daily earnings and hence the earnings risk that they face by almost 78 percent. The results demonstrate that perception bias is promising as an explanation for the lack of a risk-incentives trade-off in contractual data. Underestimating risk has a significant effect on contractual choice. We find that the greater is the degree of underestimation, the more likely workers are to select piece rates. Risk averse workers will accept risky incentive contracts with a lower earnings premium than they would if they perceived earnings risk correctly. Our evidence on the interaction between risk perception and risk preferences is inconclusive and requires further study. More than 17 percent of the workers in our sample display risk loving preferences. Yet, the interaction term between risk aversion and perception bias is not significant. Future research will need more data to measure these interaction effects.

Our results on the existence of a systematic perception bias with respect to risk are consistent with a growing literature showing that, workers have a biased perception of risk which may affect their behavior. The presence of a bias on risk has been demonstrated in different contexts, including the health risk related to smoking, alcohol and drug abuse (Viscusi, 1990; Hansen et al., 1991; Borrelli et al., 2010; Viscusi, 2016), the self-perception of weight (Rahman and Berenson, 2010), the perception of mortality risk due to different causes such as road accidents, suicide, cancer, heart diseases, stroke (Lichtenstein et al., 1978; Gegax et al., 1991; Viscusi et al., 1997, 2019) and the perception of job-related accidents (Viscusi et al., 1979; Gegax et al., 1991; Viscusi, 1993). Possible explanations to perception bias have been provided in the literature including the cognitive dissonance (Akerlof and Dickens, 1982), the cultural cognition Kahan et al. (2007) and the behavioural spillover effects (Dickinson and Oxoby, 2011). Future work may also analyze how does this perception bias influence workers productivity in risky settings.

Bibliography

Ackerberg, D. A. and M. Botticini (2002): "Endogenous matching and the empirical determinants of contract form," *Journal of Political Economy*, 110, 564–591.

Akerlof, G. A. and W. T. Dickens (1982): "The economic consequences of cognitive dissonance," *The American economic review*, 72, 307–319.

Allen, D. and D. Lueck (1992): "Contract choice in modern agriculture: cash rent versus cropshare," *The Journal of Law and Economics*, 35, 397–426.

Allen, D. W. and D. Lueck (1999): "The role of risk in contract choice," *Journal of Law, Economics, and Organization*, 15, 704–736.

Bellemare, C. and B. Shearer (2013): "Multidimensional heterogeneity and the economic importance of risk and matching: evidence from contractual data and field experiments," *The RAND Journal of Economics*, 44, 361–389.

Borrelli, B., R. B. Hayes, S. Dunsiger, and J. L. Fava (2010): "Risk perception and smoking behavior in medically ill smokers: a prospective study," *Addiction*, 105, 1100–1108.

Dickinson, D. L. and R. J. Oxoby (2011): "Cognitive dissonance, pessimism, and behavioral spillover effects," *Journal of economic Psychology*, 32, 295–306.

Dominitz, J. and C. F. Manski (1997): "Using expectations data to study subjective income expectations," *Journal of the American statistical Association*, 92, 855–867.

Eddy, A. and B. Seifert (1992): "An examination of hypotheses concerning earnings forecast errors," *Quarterly Journal of Business and Economics*, 22–37.

Gegax, D., S. Gerking, and W. Schulze (1991): "Perceived risk and the marginal value of safety," *The Review of Economics and Statistics*, 589–596.

Hansen, W. B., A. E. Raynor, and B. H. Wolkenstein (1991): "Perceived personal immunity to the consequences of drinking alcohol: The relationship between behavior and perception," *Journal of Behavioral Medicine*, 14, 205–224.

Hölmstrom, B. (1979): "Moral hazard and observability," *The Bell journal of economics*, 74–91.

Holt, C. A. and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *The American Economic Review*, 92, 1644–1655.

Kahan, D. M., D. Braman, J. Gastil, P. Slovic, and C. Mertz (2007): "Culture and identity-protective cognition: Explaining the white-male effect in risk perception," *Journal of Empirical Legal Studies*, 4, 465–505.

Lichtenstein, S., P. Slovic, B. Fischhoff, M. Layman, and B. Combs (1978): "Judged frequency of lethal events." *Journal of experimental psychology: Human learning and memory*, 4, 551.

Manski, C. F. (2004): "Measuring expectations," *Econometrica*, 72, 1329–1376.

McDonald, C. L. (1973): "An empirical examination of the reliability of published predictions of future earnings," *The Accounting Review*, 48, 502–510.

Prendergast, C. (2000): "What trade-off of risk and incentives?" *The American Economic Review*, 90, 421–425.

Rahman, M. and A. B. Berenson (2010): "Self-perception of weight and its association with weight-related behaviors in young reproductive-age women," *Obstetrics and gynecology*, 116, 1274.

Savage, I. (1993): "Demographic influences on risk perceptions," *Risk analysis*, 13, 413–420.

Shearer, B. (2004): "Piece rates, fixed wages and incentives: Evidence from a field experiment," *The Review of Economic Studies*, 71, 513–534.

Viscusi, W., J. Hakes, and A. Carlin (1997): "Measures of mortality risks," *Journal of Risk and Uncertainty*, 14, 213–233.

Viscusi, W. K. (1990): "Do smokers underestimate risks?" *Journal of political Economy*, 98, 1253–1269.

——— (1993): "The value of risks to life and health," *Journal of economic literature*, 31, 1912–1946.

——— (2016): "Risk Beliefs and Preferences for E-cigarettes," *American Journal of Health Economics*, 2, 213–240.

Viscusi, W. K., J. Huber, and J. Bell (2019): "Responsible Precautions for Uncertain Environmental Risks," *Journal of Benefit-Cost Analysis, Forthcoming*, 19–20.

Viscusi, W. K. et al. (1979): *Employment hazards: An investigation of market performance*, 148, Harvard University Press.

Appendix

Appendix 1

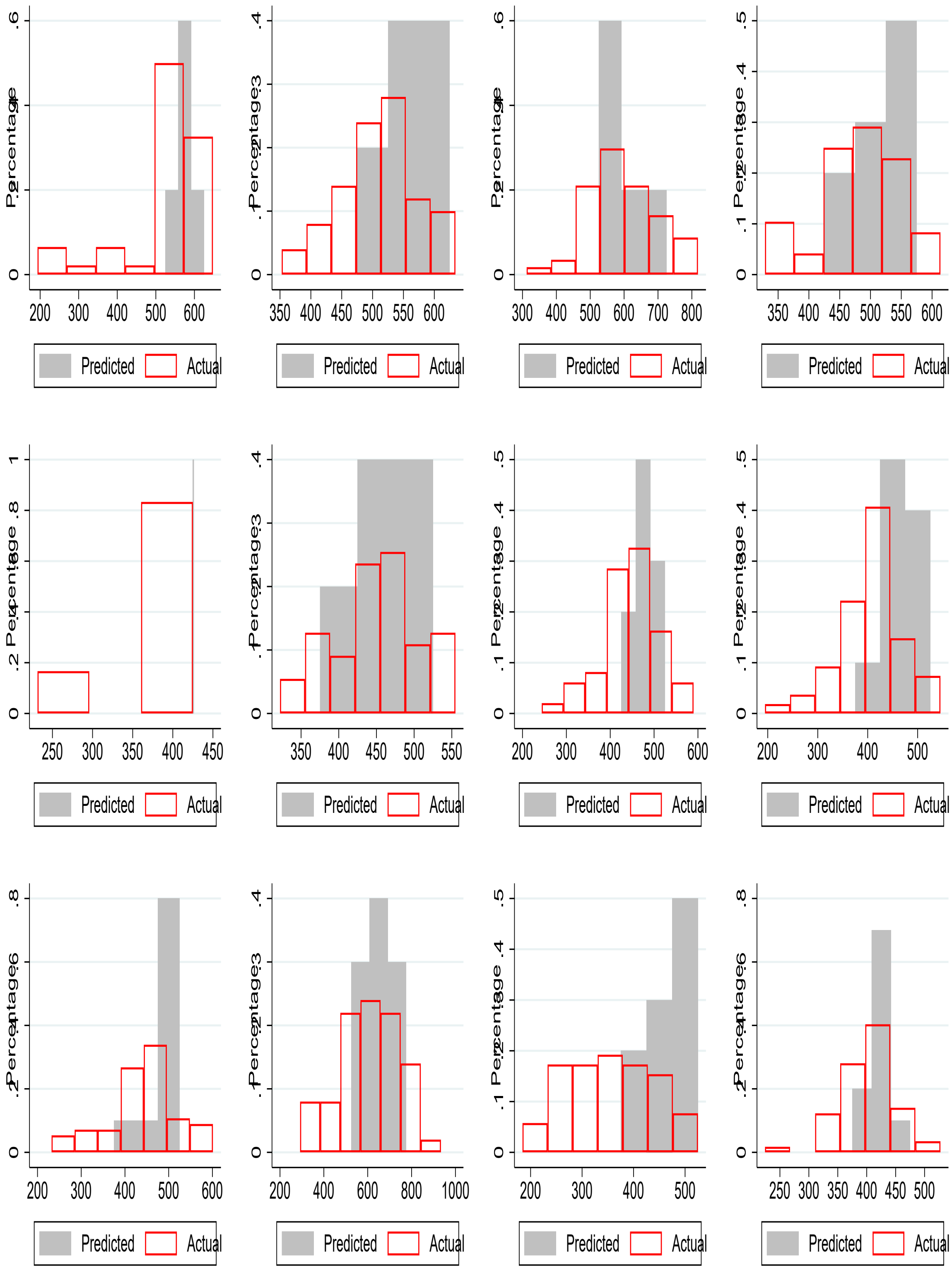
Table 2.8: Questionnaire to measure the perceived earnings

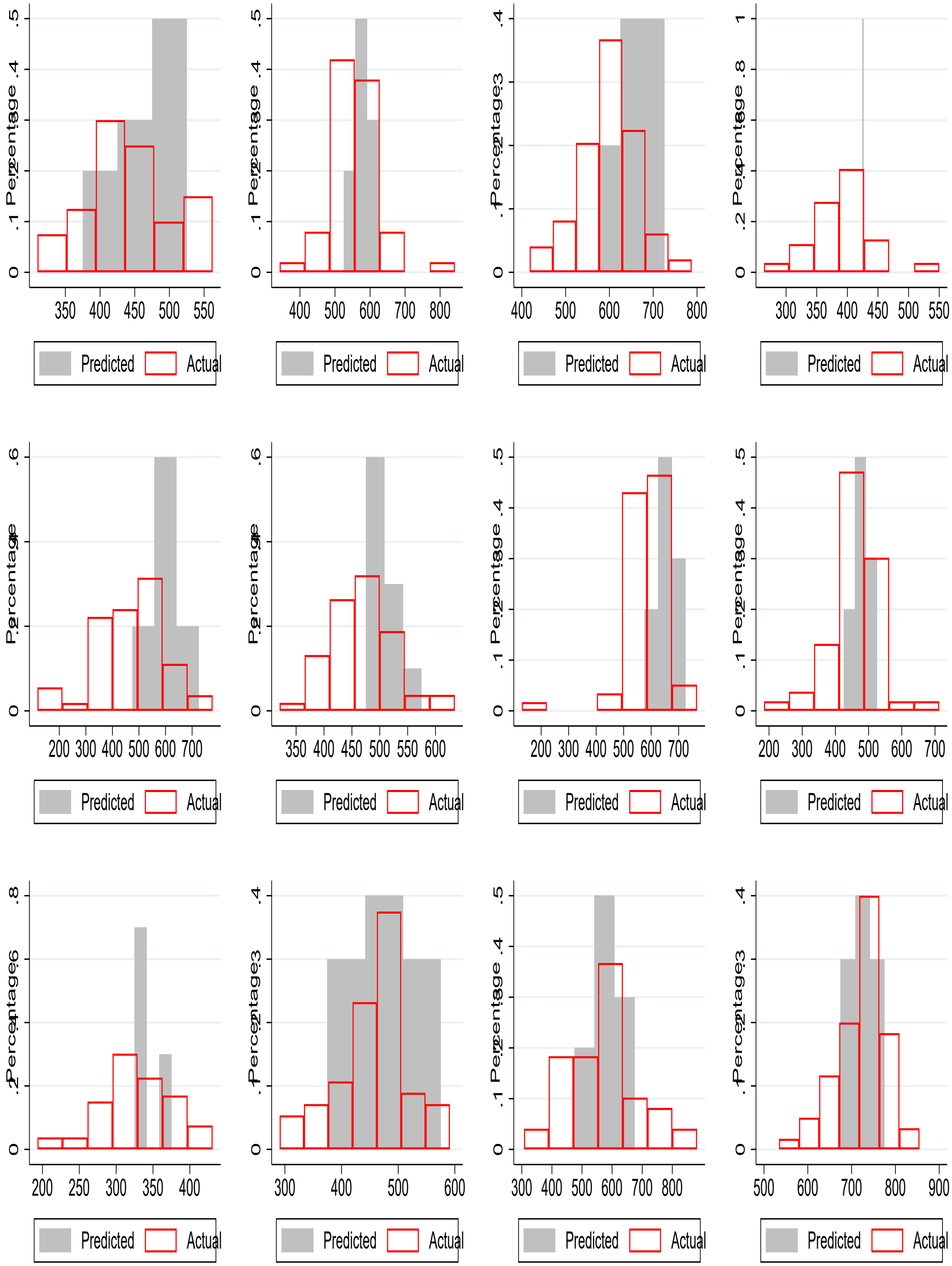
Question 1:	On a typical day I expect to earn ____ dollars
Question 2:	Over a typical 10-day period, I expect to earn:
<100	______ days out of 10
[100,150[	______ days out of 10
[150,200[	______ days out of 10
[200,250[	______ days out of 10
[250, 300[	______ days out of 10
[300, 350[	______ days out of 10
[350, 400[	______ days out of 10
[400, 450[	______ days out of 10
[450, 500[	______ days out of 10
[500, 550[	______ days out of 10
[550,600[	______ days out of 10
[600,650[	______ days out of 10
[650,700[	______ days out of 10
[700,750[	______ days out of 10
[750,800[	______ days out of 10
[800,850[	______ days out of 10
[850,900[	______ days out of 10
>900	______ days out of 10
Total	__10__ DAYS

Table 2.9: Contract choices experiment

Decision	Option A	I choose A	Option B	I choose B
1	Piece rate		Fixed wage of 100$	
2	Piece rate		Fixed wage of 150$	
3	Piece rate		Fixed wage of 200$	
4	Piece rate		Fixed wage of 250$	
5	Piece rate		Fixed wage of 300$	
6	Piece rate		Fixed wage of 350$	
7	Piece rate		Fixed wage of 400$	
8	Piece rate		Fixed wage of 450$	
9	Piece rate		Fixed wage of 500$	
10	Piece rate		Fixed wage of 550$	
11	Piece rate		Fixed wage of 600$	
12	Piece rate		Fixed wage of 650$	

Figure 2.5: Distributions of actual and average earnings for each worker




Percentage
Predicted
Actual

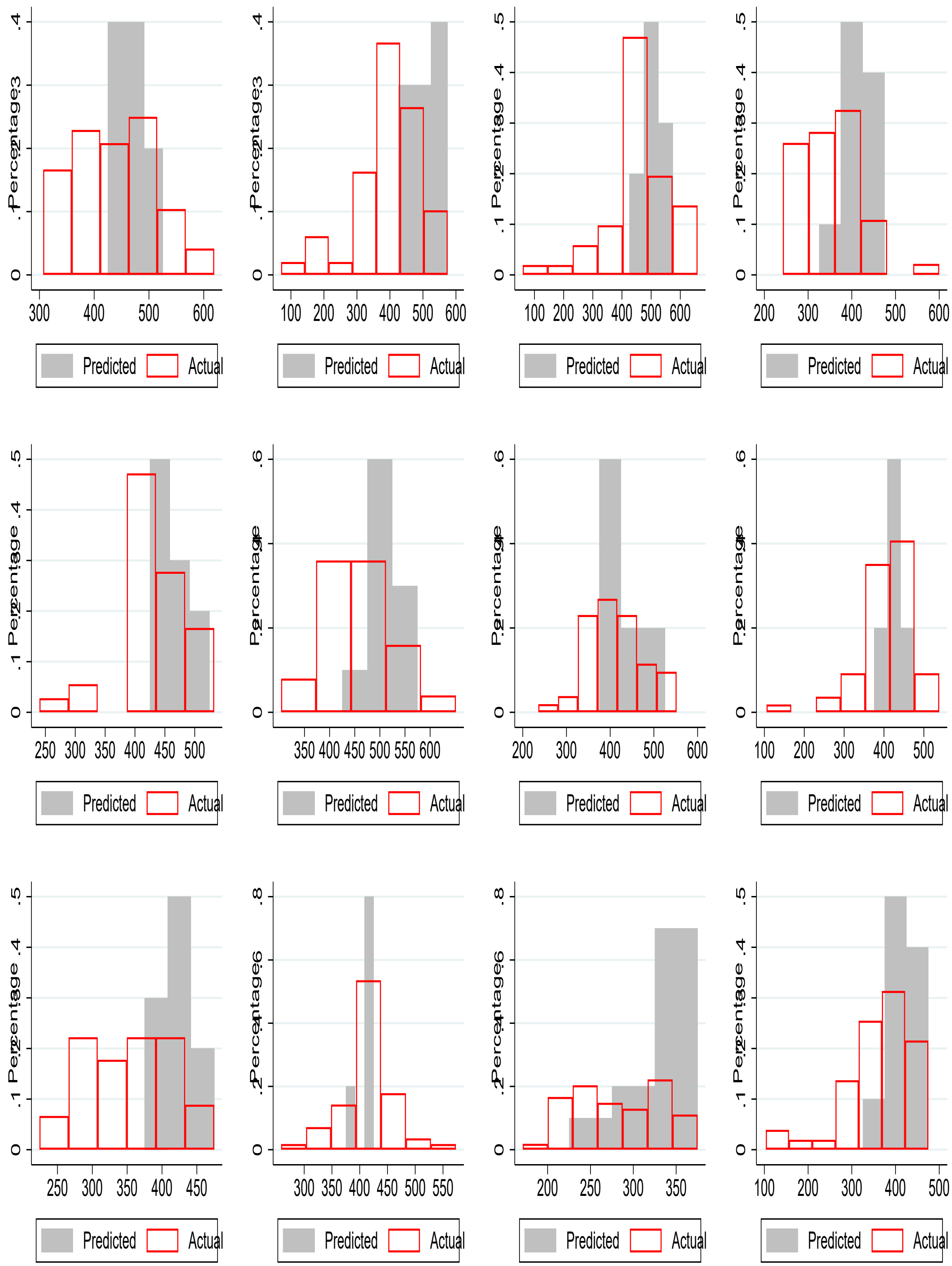

Percentage
Predicted
Actual

Table 2.10: Risk-preference-revealing-experiment

Decision	Lottery A	Lottery B	E[B]-E[A]**
1	80$(1/10), 64$(9/10)	154$(1/10),4$(9/10)	-46.80$
2	80$(2/10), 64$(8/10)	154$(2/10),4$(8/10)	-33.20$
3	80$(3/10), 64$(7/10)	154$(3/10),4$(7/10)	-20.00$
4	80$(4/10), 64$(6/10)	154$(4/10),4$(6/10)	-6.40$
5	80$(5/10), 64$(5/10)	154$(5/10),4$(5/10)	7.20$
6	80$(6/10), 64$(4/10)	154$(6/10),4$(4/10)	20.40$
7	80$(3/10), 64$(3/10)	154$(7/10),4$(3/10)	34.00$
8	80$(8/10), 64$(2/10)	154$(8/10),4$(2/10)	47.20$
9	80$(9/10), 64$(1/10)	154$(9/10),4$(1/10)	60.80$
10	80$(10/10)	154$(10/10)	74$

**choices differential between lottery A and lottery B

Table 2.11: Sensitivity analysis: Perception bias without the participant who indicates no variation in his earnings across earnings intervals

VARIABLES	Predicted average earnings Q1	Predicted average earnings Q2	Predicted Standard deviation
Absolute Bias	28.83***	43.94***	-150.25***
	(6.299)	(5.147)	(5.327)
Observations	46	46	46

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 2

We demonstrate that the risk premium is positive ($RP_{ij} > 0$) only for risk averse workers ($\delta_i < 1$).

$$\begin{aligned} RP_{ij} &> 0 \\ \bar{w}\left(\frac{k_h}{k_i}\right)^{\gamma}\left(\exp^{\frac{1}{2}(\gamma+1)^2(1-\delta_h)\sigma^2} - \exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-\delta_h)\sigma^2}\right) &> 0 \\ \frac{1}{2}(\gamma+1)^2(1-\delta_h)\sigma^2 &> \frac{1}{2}(\gamma+1)^2(\delta_i-\delta_h)\sigma^2 \\ 1-\delta_h &> \delta_i - \delta_h \\ \delta_i &< 1 \end{aligned} \tag{2.28}$$

Appendix 3

The sign of $\frac{\partial RP_{ij}}{\partial \epsilon_{ij}}$ depends on $(1-\delta_h) - (\delta_i - \delta_h)\exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-1)(\sigma^2+\epsilon_{ij})}$.

For a risk averse worker:

$$\delta_i < 1 \Longrightarrow \begin{cases} 1-\delta_h > & \delta_i - \delta_h \\ 0 < & exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-1)(\sigma^2+\epsilon_{ij})} < 1, \end{cases} \tag{2.29}$$

Equation (2.29) implies:

$$\begin{aligned} (1-\delta_h) &> (\delta_i-\delta_h)\exp^{\frac{1}{2}(\gamma+1)^2(\delta_i-1)(\sigma^2+\epsilon_{ij})} \\ \Longrightarrow \frac{\partial RP_{ij}}{\partial \epsilon_{ij}} &> 0 \end{aligned} \tag{2.30}$$

Appendix 4

We show that when we relax the assumption of Viscusi (1993); Viscusi et al. (1979), that the utility systematically decreases with risk, we find that risk preference matter in the wage-risk trade-off. For simplicity, let consider a state independent expected utility function which depends on the wage w and the risk level σ. A worker's wage-risk trade-off consists in choosing the best combination of (w, σ) which maximizes his utility. A worker's problem is:

$$\max_{(w,\sigma)} U(w,\sigma) \tag{2.31}$$

The indifference curve is given by:

$$\frac{dw}{d\sigma} = -\frac{U_\sigma}{U_w} \tag{2.32}$$

where $U_w = \dfrac{\partial U(w,\sigma)}{\partial \sigma}$ and $U_\sigma = \dfrac{\partial U(w,\sigma)}{\partial w}$.

Given that $U_w > 0$, the sign of $\frac{dw}{d\sigma}$ depends on U_σ.

- If $U_\sigma < 0$, implying that the individual is risk averse, then $\frac{dw}{d\sigma} > 0$ and the required wage rate increases with the risk level.

- If $U_\sigma = 0$, implying that the individual is risk neutral, then $\frac{dw}{d\sigma} = 0$ and the required wage rate does not depend on the risk level.

- If $U_\sigma > 0$, implying that the individual is risk loving, then $\frac{dw}{d\sigma} < 0$ and the required wage decreases with the risk level.

Chapter 3

Motherhood and Female Self-Employment: Theory and Evidence from Uganda

3.1 Résumé

La présence d'enfants en bas âge dans le ménage est bien souvent indexée dans la littérature comme le déterminant principal de l'emploi des femmes. Toutefois, peu d'évidences empiriques existent dans les pays en développement pour confirmer cette idée. Dans cet article, nous examinons l'impact de la maternité et de l'éducation sur l'emploi autonome des femmes en Ouganda. Notre stratégie empirique consiste à corriger le biais de sélection lié à la participation des femmes sur le marché du travail et l'endogénéité de la maternité et de l'éducation. Les résultats suggèrent que la présence d'enfants en bas âge n'a pas d'impact significatif sur la probabilité d'une femme d'exercer un emploi autonome en Ouganda. A l'inverse, l'éducation a un effet causal négatif et significatif sur l'emploi autonome des femmes. Ce qui implique qu'une meilleure éducation accroît la probabilité pour une femme d'accéder à un emploi salarié sur le marché du travail. En termes de politiques, nos résultats impliquent que l'amélioration de l'éducation des filles est une politique plus efficace pour permettre aux femmes d'accéder à des emplois salariés que les politiques visant la réduction du nombre d'enfants dans la pays en développement.

Mots-clés: Maternité, Education, Femme, Emploi autonome, Endogeneité, Biais de sélection

Codes JEL : H52, I26, J21, J24, J46

3.2 Abstract

The leading explanation of women's self-employment in the labor literature is motherhood but supporting evidence exists only for developed countries. Indeed, in developing countries where women's self-employment dwarf international standards and is predominantly a feature of the insecure informal sector, little is known about its causes. This paper investigates the importance of motherhood to women's self-employment. Using micro-level data from Uganda, we estimate the effects of motherhood and education on the self-employment probabilities of women in Uganda. Our estimation framework accounts for selection bias and the endogeneity of motherhood and education jointly. Consistent estimators of the effects of motherhood and education are obtained by estimating a four-equation model, including self-employment, labor force participation, education and motherhood equations. We find no evidence of a causal effect of motherhood on women's self-employment probabilities. In contrast, education has a negative causal effect on these probabilities. Both these results differ from the existing literature showing that motherhood, not education, drives women's self-employment probabilities in developed countries. Our results suggest that in the developing world, public policies most effective at reducing the gender gap in pay are likely those that pull women out of self-employment, through better education and access to affordable childcare services.

Keywords: Motherhood, Education, Women, Self-Employment, Endogeneity, Sample selection bias

Codes JEL : H52, I26, J21, J24, J46

3.3 Introduction

The literature focusing on women's employment outcomes often points to motherhood not only as a barrier to their participation in the paid labor force (Agüero and Marks, 2011), but also as the driver of their self-employment probabilities (Connelly, 1992; Semykina, 2018). Self-employment is presented as an occupational strategy allowing women to balance work and family commitments (Connelly, 1992; Craig et al., 2012). However, these findings are limited to developed countries, where women's self-employment probabilities are lower than those for men, and are much smaller than for developing countries.[1] Moreover, these findings are also being obtained with empirical strategies that overlook the interrelation of education and self-employment decisions.

In this paper, we present new quantitative evidence on the importance of motherhood to women's self-employment in a developing economy's context. The first contribution of this paper is to explore theoretical answers to the following question: does motherhood matter to women's self-employment in the context of a developing economy? We address this issue using a simple Becker-type unitary model of women's employment decisions. In this model economy, women either participate in the paid labor force, or not. Those who do not participate specialize their time to caring for their younger children, if any. Those who participate can choose between wage employment and self-employment. Wage employment takes place in the formal sector, and self-employment, in the informal sector (Chen, 2001; Heintz and Valodia, 2008; Ulrichs, 2016) characterized by insecure property rights (Heintz and Valodia, 2008). Consistent with empirical evidence, in any household, caregiving to younger children is primarily a maternal responsibility (Wodon and Blackden, 2006; Budlender, 2008; Heath, 2017; Doss, 2018; Semykina, 2018). As children are not allowed in a formal sector workplace, motherhood raises the opportunity cost of being wage-employed for two related reasons. First, nonparental caregiving adds to household costs, through either household purchase of it in the market, or the living expenses of a live-in nanny, as is often the case in a developing country. The second potential reason is the imperfect substitutability between parental and nonparental caregiving. As parental interactions account for children's cognitive and socioemotional development, if not adequately compensated for by nonparental caregiving, employment-induced restrictions on mother's interactions with her younger children may impede their development. These two reasons comprise the motherhood penalty for wage employment. However, in the context of a developing economy, self-employment also has two contributing sources of motherhood penalty. First, unlike in a developed economy, self-employment in a developing economy takes place predominantly in the insecure informal sector, where labor is less productive than in the formal sector (Heath, 2017; Heintz and Valodia, 2008). Second, in developing countries, self-employed women often care for their younger children while working, and this too reduces labor productivity in

[1] The United States of America are a case in point. Indeed, in 2012, the female-to-male ratio in self-employment was 35.7 percent, compared to 48.6 percent in wage and salary employments; while women had on average a self-employment probability of only 7.1 percent, compared to 11.5 for men (Roche, 2014)

self-employment (Samman et al., 2016), in addition to potentially disrupting the mother's interactions with her younger children, either due to fatigue or lack of specialization of her time to caregiving. Therefore, since both wage employment and self-employment entail motherhood penalties, relevantly for empirical investigation, the model shows that the effect of younger children on women's self-employment is ambiguous. In other words, the issue of whether motherhood matters to women's self-employment does not have a clear theoretical answer in a developing economy context.

The second contribution of this paper is to investigate the importance of motherhood to women's self-employment empirically, using micro-level data from the Uganda 2016 Demographic and Health Surveys (DHS). The challenge of identifying the causal effect of motherhood on women's self-employment probabilities stems from the potential presence of three problems. Semykina (2018) highlights the nonrandom self-selection of working-age women into the paid labor force and the endogeneity of fertility decisions as two problems that can complicate the analysis of the causal effect of motherhood on women's self-employment probabilities. But there is a third one. It stems from the need to account for the confounding effect of education on the effect of motherhood. However, controlling for female education is complicated by the potential joint determination of women's schooling and employment outcomes (Ahituv and Tienda, 2004), of which self-employment is an integral part. Indeed, in a developing country's setting where norms and traditions play an important role in women's lives, unobservables that influence both women's education outcomes and their self-employment probabilities are very likely to be present. For example, gendered norms that lead parents to under-invest in their daughters' education are also likely to influence these daughters' career paths later when they marry and have children. Not accounting for this potential joint determination can lead to an underestimation of the effect of education on women's self-employment probabilities.

We address all these three problems jointly. To correct for the endogeneity of the motherhood variable, we follow Aguero and Marks (2008); Agüero and Marks (2011) in instrumenting fertility with infertility shocks. To correct for the endogeneity of education, we exploit a policy reform that established universal primary education (UPE) nationwide in Uganda. This policy reform removed all school fees at the primary school level for all primary school-age children from 1997 on (Deininger, 2003; Grogan, 2008). Based on their year of birth, women in our sample had either no exposure, partial exposure, or full exposure to the UPE reform. This fact makes a binary conceptualization of the instrumental variable for education inappropriate, as it would either live out women who had partial exposure to the UPE reform or, if included, would overestimate the effect of the UPE reform on women's educational attainment. To address this additional challenge to identification, we model the UPE-based instrumental variable for education as a count variable. Consistent estimates are obtained by implementing a Full Information Maximum Likelihood Estimator method combined with a control function approach for endogenous regressors.

When we account for sample selection and the endogeneity of multiple regressors jointly, we find no evidence of a causal effect of motherhood on women's self-employment probabilities in a developing economy's context. This result indicates that the reported positive effect of motherhood on women's self-employment probabilities in developed economies (Connelly, 1992; Semykina, 2018) is not universal: it does not extend to the context of a developing economy, where women's self-employment is predominantly a phenomenon of the insecure informal sector, making it a low-productivity economic activity. Importantly, we find evidence that low levels of education push working women into self-employment as a source of livelihood: an additional year of schooling completed reduces a woman's probability of self-employment by 3.05 percentage points. This figure is roughly three times larger when compared to an estimation approach where we do not correct for any potential identification problem mentioned above.

Overall, the main takeaway from our theoretical model and empirical findings is that the context of women's self-employment determines the nature of social factors that drive its incidence. In developed countries, women's self-employment takes place predominantly in the formal sector where women can take advantage of the rule of law to hire labor and enforce contracts. In this context, motherhood is more likely to pull women into self-employment as a career path that allows them to balance between paid work and their family responsibilities. By contrast, in developing countries, where self-employment is predominantly a phenomenon of the informal sector characterized by insecure property rights, women's involvement in this type of activity is likely to be highly involuntary, given the risks involved and the resulting restrictions imposed on their spatial mobility in response to these risks. In this context, women's self-employment is more likely to be driven by push factors, such as the inability to find a wage job in the formal sector, due to low levels of education.

Our study contributes to the literature relating women's individual characteristics to their employment outcomes (Aguero and Marks, 2008; Agüero and Marks, 2011; Agüero et al., 2012; Heath, 2017; Semykina, 2018). Aguero and Marks (2008) use infertility shocks as a source of an exogenous variation in fertility outcomes, to analyze the causal effect of motherhood on women's labor force participation, in the context of Latin America. They find no evidence that children have a causal effect on women's labor force participation. Their definition of labor force participation includes participation in both paid and unpaid labor. Agüero and Marks (2011) also exploit infertility shocks to investigate the causal effect of motherhood on women's employment outcomes, in the context of 26 developing countries. They find that the presence of children limits women's chances of engaging in paid employment. Our paper differs from Aguero and Marks (2008) and Agüero and Marks (2011) by going beyond women's labor force participation decisions, to include social factors underlying women's self-employment probabilities, while also accounting for selection bias as well as the endogeneity of women's education levels.

Other papers investigating the causal effect of motherhood on women's self-employment probabilities are Connelly (1992), Bianchi (2000), Wellington (2006), Boserup et al. (2013), Lim (2018). In addition, Semykina (2018) takes account of sample selection and endogeneity. Yet all these papers treat data from developed countries. For example, Semykina (2018) finds that, in a developed economy, the presence of younger children increases women's self-employment probabilities, while education has no effect on these probabilities. Our study which focuses on the context of a developing economy draws attention to the fact that these findings are not universal, and suggests that the sociocultural context in which self-employment takes place determines the nature of factors that drive women's choice of this career path. This has implications for public policies likely to reduce the gender pay gap in developing countries.

The remainder of this paper is structured as follows. Section 2 presents the theoretical model from which we derive some testable predictions. Section 3 examines the data and presents the stylized facts while Section 4 presents the empirical model. Section 5 reports the empirical results of our estimation. Finally, Section 6 discusses the results and concludes.

3.4 Theoretical Model

We seek to understand the manner in which female characteristics influence their employment outcomes, distinguishing between those characteristics that influence their decisions to participate in the labor force, and those characteristics that influence the sector of employment (self-employment or wage employment).

The main features of our model draw from a number of existing household models of occupational choice, including Fernández et al. (2004), De la Croix and Doepke (2004), Tertilt (2006) and Bertrand et al. (2016). We adapt these existing models to fit the context of a developing economy, and generate predictions about the effects of individual characteristics on women's self-employment.

3.4.1 Fundamentals

There is a measure one of households indexed by $i \in [0, 1]$. A typical household in this environment is identified by a vector containing the woman's individual characteristics $X_i = \left(E_i, K_5^i, M_i\right)$, where $E_i \geq 0$ denotes her level of education, $M_i \in \{0, 1\}$, her marital status, and $K_5^i \geq 0$, the number of preschool children she has.

We classify households into two types: single-parent households, headed by a woman ($M_i = 0$), and two-parent households, co-headed by a man and a woman ($M_i = 1$). These are the dominant family types observed in the developing world (*World Family Map 2015*)[2]. We also take the intra-household

[2]See the WORLD FAMILY MAP 2015. *Mapping Family Change and Child Well-Being Outcomes*. Available

division of labor as exogenously given.

An important feature of our model is that caregiving is the sole responsibility of the mother. This internal household organization mirrors the intra-household gendered division of labor prevalent in developing countries, particularly those from sub-Saharan Africa, where women spend more than three times as much time as men on unpaid domestic work, including childcare and home-making (Budlender, 2008). Working mothers with preschool children thus face a time-constraint for their participation in the labor force, inducing dual-earning households (those in which the woman has marital status $M_i = 1$) to choose between a traditional single-earner family model ($d_i = 0$) and a dual-earner family model ($d_i = 1$) to support the family. In other words a married woman's labor force participation decision is intertwined with the household choice of family model. There is no involuntary unemployment among women in this environment. However, a woman must incur a search cost, $\underline{c} > 0$, to secure a wage employment. There is no such cost for self-employment.

In a co-headed household, the male spouse's income denoted as θ_i is exogenously given, and thus is a contributing source of heterogeneity across women. In a co-headed traditional single-earner household ($d_i = 0$), the woman specializes completely in family affairs (caregiving and homemaking). In a dual-earner household ($d_i = 1$), the woman is either self-employed ($y_i = 1$) or wage-employed ($y_i = 0$). The decision on the woman's occupational profile, $(d_i, y_i) \in \{0, 1\} \times \{0, 1\}$, is made by the household as a unit, to maximize an appropriately defined household utility function.

Finally, we assume that a woman's level of education, E_i, and the number of preschool children she has, K_5^i, are exogenously given. This assumption allows us to focus on deriving the causal effects these two characteristics may have on women's employment outcomes. Furthermore, we restrict attention to the presence of preschool children in the household, ignoring school-age children. This emphasizes the need to care for preschool children —either directly by the mother, or through childcare professionals—, which introduces a constraint on a mother's occupational choice. Evidence shows that many developing countries have near universal primary school enrollment (UNDP's *Human Development Report 2016*), implying that mothers of school-age children are relatively free to participate in the labor force.

A single female has no choice but to participate in the paid labor force, to support herself and her children, if any. Her only decision, therefore, concerns her self-employment status $y_i \in \{0, 1\}$.

3.4.2 Occupation-Specific Time Allocation Constraint

Suppose household i chooses the occupational profile $(d_i, y_i) = (0, y_i)$. In this case, household i is a family in which the male works for pay and the female prioritizes care-giving and homemaking over

online at https://ifstudies.org/ifs-admin/resources/reports/wfm-2015-forweb.pdf paid employment, and so does not participate in the paid labor force. She earns no income and thus is totally dependent on her husband for her needs and those of the children they have, if any.

Next, suppose household i chooses the occupational profile $(d_i, y_i) = (1, 1)$. Then, the woman participates in the paid labor force ($d_i = 1$) as a self-employed worker ($y_i = 1$) and takes care of her children while working. This type of work arrangement for self-employed mothers is very common in developing countries (Quisumbing et al., 2007; Mousié, 2016). As we show below, this balancing act has consequences for a woman's productivity in paid work.

Finally, suppose household i chooses the occupational profile $(d_i, y_i) = (1, 1)$ for the woman. Then, the woman fully prioritizes paid employment ($d_i = 1$) over family responsibilities, with participation in wage employment ($y_i = 1$). Wage employment takes place exclusively in the formal sector, and any woman can secure a wage job provided she pays the search cost, $\underline{c}$. In this sector, no children are allowed in the workplace so households with wage-employed mothers must pay for childcare services from professionals caregivers during the time they are at their paid job.[3]

3.4.3 Preferences and Budget Constraints

Each household derives utility from the spouses' joint consumption of a numeraire, c_a^i, from a household public good, denoted as G_i.[4] A household utility payoff writes as follows:

$$u_i = \log c_i^a + \gamma \log G_i, \tag{3.1}$$

where $\gamma > 0$.

The level of household i's public good, G_i, is a composite of the quantity and quality of children (Becker and Lewis, 1973; Becker and Tomes, 1976). We follow the literature in assuming that the production technology for the household public good is Cobb-Douglas in the quantity and quality of children (De La Croix and Doepke, 2003; De la Croix and Doepke, 2004).[5] A household i with K_5^i preschool children nurtured into achieving each a level of quality, q_i, generates a level of household public good given by:

$$G_i = (q_i)^\beta (K_5^i)^{1-\beta}, \tag{3.2}$$

where $\beta \in (0, 1)$. In Equation (3.2), by quality of offspring we mean the level of cognitive and socioemotional development attained by a child. Consistent with the literature on child development (Becker and Tomes, 1976; Currie and Moretti, 2003; Del Boca et al., 2013; Heckman and Mosso, 2014; Heath, 2017), we model child quality as produced through a Cobb-Douglas function of monetary and

[3]In most developing countries, nurseries and daycare centers are rare. Live-in nannies or domestic workers are the main form of professional childcare providers (Hein and Cassirer, 2010)

[4]In a household headed by a single woman, c_i^a denotes her own consumption.

[5]We ignore school-age or grown up children for simplicity.

effective time inputs. On one hand, household income is necessary to meet the nutritional needs of children. On the other hand, caregiving to younger children plays an important role in their cognitive and socioemotional development, with long-term implications for their human capital outcomes. Recall that in our model environment, caregiving is primarily a maternal, and thus must be out-sourced to non-parental childcare providers when the mother is wage-employed. We assume that the effect of caregiving on child quality, q_i, depends not only on the time spent interacting with children, but also on the quality of those interactions, as measured by the mother's level of education, E_i. Therefore, in a single-earner household—i.e., one in which the mother specializes completely in caregiving to her younger children—, child quality is given by:

$$q_i = (1 + E_i)^{1-\eta} \left(c_i^k\right)^{\eta}, \tag{3.3}$$

where $\eta \in (0, 1)$. Child quality increases with mother's education because maternal education impacts the quality of the interactions a mother has with her younger children, which, in turn, affects their cognitive and socioemotional development.

A self-employed mother (i.e., $y_i = 1$) interacts with her younger children while working. Arguably, this lack of specialization may adversely impact the transmission of human capital from mother to children, due to either fatigue or preoccupations with the outcome of business operations. For a self-employed mother of K_5^i preschool children, therefore, child quality is given by:

$$q_i = (1 + \delta_1\ E_i)^{1-\eta} \left(c_i^k\right)^{\eta}, \tag{3.4}$$

where $\delta_1 \in [0, 1]$ measures the impact lack of specialization of mother's time to caregiving has on the child's cognitive and socioemotional development. As long as $\delta_1 < 1$, lack of maternal specialization in caregiving to her younger children adversely impacts their cognitive and socioemotional development. By contrast, $\delta_1 = 1$ means there is no such adverse effect.

A wage-employed mother (i.e., $y_i = 0$), has limited interactions with her younger children because she works in the formal sector where children are not allowed in the workplace. Children of wage-employed mothers thus receive nonparental care from professionals. In the literature on child development, there are concerns that a child's separation from the primary caregiver— the mother—, may trigger stress and anxiety in the child, with adverse effects on his or her socioemotional development (Mercer, 2006). Echoing these concerns, we model the impact of a wage-employed mother's limited interactions with her younger children as measured by $\delta_0 \in [0, 1]$. Consequently, child quality for a wage-employed mother is given by:

$$q_i = (1 + \delta_0\ E_i)^{1-\eta} \left(c_i^k\right)^{\eta}, \tag{3.5}$$

When $\delta_0 = \delta_1 = 1$, parental and nonparental caregiving are perfect substitutes in the production of child quality. Unless

$$\delta_0 < \delta_1 \leq 1, \tag{3.6}$$

there is no balancing act motive for women's self-employment. It is important to note, however, that the literature addressing parental versus nonparental caregiving in relation to child development yields no consensus on the ranking of these caregiving strategies. Some studies find that nonparental childcare arrangements adversely impact child development when the quality of such care is poor (Baker et al., 2008; Herbst, 2013; Fort et al., 2017), while others find that it boosts child development where the quality is high (Gupta and Simonsen, 2010; Noboa-Hidalgo and Urzua, 2012; Drange and Havnes, 2014). Moreover, with respect to the role of parental interactions for children's long-term development, several recent studies find no effect (Rasmussen, 2010; Dustmann and Schönberg, 2012; Baker and Milligan, 2015), while several others find some positive effects on child human capital (Carneiro et al., 2015; Danzer and Lavy, 2018; Drange and Havnes, 2019). We discuss our theoretical results in light of this lack of consensus in the literature on parental vs. nonparental caregiving to preschool children.

As in Fernández et al. (2004) and Bertrand et al. (2016), in a dual-earner household, spouses pool their individual incomes. We denote as $R(d_i, y_i, X_i)$ total household i income, where $X_i = \left(E_i, K_5^i, M_i\right)$ denotes the vector of the woman's individual characteristics, for all $i \in [0,1]$. We denote a woman's occupation-specific income as $W(y_i, X_i)$. A household i's occupation-specific income thus writes as follows:

$$R(d_i, y_i, X_i) = \theta_i M_i + d_i[W(y_i, X_i) - (1 - y_i)\underline{c}], \tag{3.7}$$

for all (d_i, y_i, X_i), where θ_i denotes the male partner's exogenously given income. The structure of expression (3.7) reflects the fact that a household that chooses the occupation profile $(d_i, y_i) = (1, 0)$ faces an additional expenditure in the form of a search cost to secure a wage employment.

For a working woman $(d_i = 1)$ with individual characteristics X_i, income earned depends on the type of employment she holds, y_i. The income of a self-employed woman $(y_i = 1)$ is generated from the operation of a constant-return-to scale technology that requires effective labor only, which is the product of a woman's raw labor time, in total quantity normalized to unity, and her level of education, E_i. Such woman cares for her preschool children while doing paid work. We assume that an hour of work in self-employment yields an income $\underline{A}E_i$ for a self-employed woman with no preschool children, where $\underline{A} > 0$ is a productivity factor. This modeling assumption is consistent with evidence showing that even in a developing economy's informal sector, the return to self-employment increases with the worker's level of education (Moock et al., 1989). In the presence of $K_5^i > 0$ preschool children, the same hour of work yields an income $\underline{A}(1 - \nu K_5^i)E_i$, because caring for children while working induces loss of effective work time, which in turn induces loss of income (Samman et al., 2016). Indeed, the difference,$\underline{A}E_i - \underline{A}\left(1 - \nu K_5^i\right)E_i = \nu \underline{A} E_i K_5^i$ is the forgone income of balancing motherhood and paid work in self-employment. We refer to this cost as the motherhood penalty in self-employment. The labor income of a self-employed woman with a vector of individual characteristics X_i thus is

$W(1, X_i) = \underline{A}\left(1 - \nu K_5^i\right) E_i$.

Assumption 1. $\nu \bar{K} < 1$, *where* $\bar{K} = \max_i K_5^i$

Assumption 1 states that the forgone income of balancing motherhood and paid work in self-employment is not too high, in the sense that it does not preclude participation in this type of activity.

In wage employment, a woman also allocates her entire unit endowment of time to paid work and provides no care-labor. During the entire time she is at work, her younger children receive nonparental care by a professional. We assume that the production technology in wage employment is also constant-return-to scale to effective labor. An hour of work in wage employment yields an income equal to AE_i. This modeling is consistent with empirical evidence showing that the return to wage employment rises with a worker's level of education (Garcia-Mainar and Montuenga-Gomez, 2005).

Assumption 2. $A > \underline{A}$

Assumption 2 states that effective labor is more productive in wage employment than in self-employment. In other words, there is a productivity premium for wage employment, and this premium rises with a woman's level of education, E_i. This assumption can be justified by the fact that, in developing countries, self-employment is a phenomenon of the informal economy, where property rights are insecure (Brown, 2015). Such insecurity acts as an impediment to the adoption of more productive technologies in self-employment (Fu et al., 2018).

Like her self-employed counterpart, a wage-employed woman allocates her entire unit endowment of time to work. Her labor income thus is given by $W(0, X_i) = AE_i$. More generally, we can rewrites the employment-specific income of a female worker as follows:

$$W(y_i, X_i) = \begin{cases} AE_i & \text{if } y_i = 0 \\ \\ \underline{A}\left(1 - \nu_i K_5^i\right) E_i & \text{if } y_i = 1 \end{cases} \tag{3.8}$$

for all i such that $d_i = 1$.

The budget constraint faced by a household with a voluntarily unemployed woman ($d_i = 0$) is given by $c_i^a + c_i^k K_5^i \leq R(0, y_i, X_i) \equiv \theta_i$, and that of a household with a self-employed woman is $c_i^a + c_i^k K_5^i \leq R(1, 1, X_i)$.

For a household with K_5^i preschool children in which the woman is wage-employed, if we denote as p_c the cost of an hour of childcare, then total childcare expenditures amount to $p_c K_5^i$. As a result, the budget constraint of a household with a wage-employed woman is $c_i^a + c_i^k K_5^i + p_c K_5^i \leq R(1, 1, X_i)$, due to childcare costs. More generally, the budget constraint of household i with a woman X_i whose occupational profile is $(d_i, y_i) \in \{0,1\} \times \{0,1\}$ obtains as follows:

$$c_i^a + c_i^k K_5^i + (1 - y_i)\, d_i p_c K_5^i \leq R\,(d_i, y_i, X_i)\,, \tag{3.9}$$

for all i.

Each household i has three sequential decisions to make, given X_i. First, the household decides on the woman's labor force participation $d_i \in \{0,1\}$. This decision is akin to choosing a family model (either a single-earner, or a dual-earner, model). A co-headed household in which the woman does not participate in the paid labor force ($d_i = 0$) is one that opted for the single-earner model. For this type of household, the next decision is on how to allocate its total income across household expenditures. By contrast, a dual-earner household ($d_i = 1$) must next chooses the type of employment for the female spouse: self-employment ($y_i = 1$) or wage employment ($y_i = 0$). Based on the outcome of this employment type decision, household i then decides on the allocation of its total income across household expenditures.

3.4.4 Occupation-Specific Payoffs

We next characterize household i's utility payoff associated with each occupational profile $(d_i, y_i) \in \{0,1\} \times \{0,1\}$, for the woman. We denote as $U^i\,(d_i, y_i, X_i)$ the utility payoff of a household in which the woman has an occupational profile (d_i, y_i), when her individual characteristics are given by X_i.

The Non Participation Payoff

An unemployed woman has no income of her own. Therefore, the decision not to participate in the paid labor force is only relevant for married women. Consequently, we assume that all women not involved in paid work are married ($M_i = 1$), and can be supported by their spouses.

From (3.1), substituting in (3.2), (3.3), (3.7), and (3.9), yields the utility payoff of a single-earner, co-headed, household ($d_i = 0$), as follows:

$$U^i\,(0, y_i, X_i) = \log\left(\theta_i - c_i^k K_5^i\right) + \eta\gamma\beta \log\left(c_i^k\right) + (1-\eta)\,\gamma\beta \log\,(1 + E_i) + \gamma\,(1-\beta) \log K_5^i, \tag{3.10}$$

for all i, such that $(d_i, y_i) = (0, y_i)$. For this household, the problem is to choose c_i^k to solve the following problem:

$$\max_{c_i^k} U^i\,(0, y_i, X_i)\,.$$

The interior solution to this problem is

$$c_i^k = \frac{\eta\gamma\beta\theta_i}{(1+\eta\gamma\beta)\,K_5^i} \tag{3.11}$$

$$c_i^a = \frac{\theta_i}{1+\eta\gamma\beta} \tag{3.12}$$

for all i such that $(d_i, y_i) = (0, y_i)$. Expression (3.11) illustrates the well-known quantity-quality trade-off (Becker and Lewis, 1973), whereby having more children (i.e., a higher K_5^i) reduces the quality each child can attain.

Substituting (3.11) and (3.12) back into (3.10), rearranging terms, yields the optimal utility payoff of non participation in the paid labor force as follows:

$$U^{i*}(0, y_i, X_i) = (1 + \gamma\beta\eta)\log\theta_i + (1-\eta)\gamma\beta\log(1+E_i) + \gamma[1-\beta(1+\gamma\eta)]\log K_5^i + Z, \tag{3.13}$$

where

$$Z_0 := \gamma\eta\beta\log\gamma\eta\beta - (1+\gamma\eta\beta)\log(1+\gamma\eta\beta).$$

Observe from expression (3.13) that unless households in this environment have a preference bias towards quantity, relative to quality, of offspring, i.e.,

$$\beta < 1/(1+\gamma\eta), \tag{3.14}$$

having children would reduce a household well-being. Furthermore, the payoff of non participation rises with the female spouse's level of education: $\partial U^{i*}(0, y_i, X_i)/\partial E_i > 0$. This is because a mother's level of education is a contributing factor to the quality of children she has. Furthermore, if condition (3.14) holds, then motherhood (i.e., having children) increases the utility payoff of non participation in the paid labor force: $\partial U^{i*}(0, y_i, X_i)/\partial K_5^i > 0$. This effect stems simply from the fact that as a unit, spouses have a preference bias towards quantity relative to quality of offspring.

The Self-Employment Payoff

Next, we characterize the utility payoff of a household in which the woman is self-employed. Suppose household i chooses the play of the sequential strategy profile $(d_i, y_i) = (1, 1)$. According to this profile, a woman (married or not) participates in the paid labor force ($d_i = 1$), earning an income as self-employed worker ($y_i = 1$). Using (3.7) and (3.8), yields this household's total income as follows:

$$R(1, 1, X_i) = \theta_i M_i + \underline{A}\left(1 - \nu K_5^i\right)E_i. \tag{3.15}$$

From (3.1), substituting in (3.2), (3.3), (3.9), and (3.15), yields the utility payoff of this household as follows:

$$\begin{aligned} U^i(1, 1, X_i) = \log\left[\theta_i M_i + \underline{A}E_i - \nu\underline{A}K_5^i E_i - c_i^k K_5^i\right] + \gamma\beta\eta\log c_i^k \\ + (1-\eta)\gamma\beta\log(1+\delta_1\ E_i) + \gamma(1-\beta)\log K_5^i, \end{aligned} \tag{3.16}$$

for all i such that $(d_i, y_i) = (1, 1)$. A typical household i for which the woman is self-employed thus solves the following problem:

$$\max_{\langle c_i^k \rangle} U^i(1, 1, X_i),$$

for all i such that $(d_i, y_i) = (1,1)$. The interior solution to this problem writes as follows:

$$c_i^a = \frac{1}{1+\gamma\beta\eta}\left[\theta_i M_i + \underline{A}E_i - \nu \underline{A} K_5^i E_i\right] \tag{3.17}$$

$$c_i^k = \frac{\gamma\beta\eta}{1+\gamma\beta\eta}\left[\frac{\theta_i M_i + \underline{A}E_i - \nu \underline{A} K_5^i E_i}{K_5^i}\right] \tag{3.18}$$

From (3.16), substituting in (3.17) and (3.18), rearranging terms, yields the optimal payoff of self-employment as follows:

$$\begin{aligned} U^{i*}(1,1,X_i) = {} & (1+\gamma\beta\eta)\log\left(\theta_i M_i + \underline{A}E_i - \nu \underline{A} K_5^i E_i\right) \\ & + (1-\eta)\gamma\beta \log(1+\delta_1\ E_i) + \gamma\left[1-\beta(1+\gamma\eta)\right]\log K_5^i + Z, \end{aligned} \tag{3.19}$$

Observe from (3.19) that

$$U^{i*}(1,1,X_i)\mid_{M_i=1} - U^{i*}(1,1,X_i)\mid_{M_i=0} > 0,$$

implying that the payoff of self-employment is higher for a married, than for an unmarried, woman. Furthermore, partial differentiation of (3.19) shows that $\partial U^{i*}(1,1,X_i)/\partial E_i > 0$. Being more educated increases the payoff of self-employment. This is because education increases the productivity of a self-employed woman. By contrast, under condition (3.14),

$$\frac{\partial U^{i*}(1,1,X_i)}{\partial K_5^i} = \frac{\gamma\left[1-\beta(1+\gamma\eta)\right]}{K_5^i} - \frac{\nu(1+\gamma\beta\eta)\underline{A}K_5^i}{\theta_i M_i + \underline{A}E_i - \nu\underline{A}E_i K_5^i}$$

has an ambiguous sign. This ambiguity stems from the two opposite effects of the presence of preschool children. Because a woman in our model derives utility from having children, their presence has a positive effect on the payoff of self-employment. However, because caring for children reduces a the productivity of a self-employed woman, the presence of children has a negative effect on this payoff through this channel.

The Wage Employment Payoff

A household in which the female spouse is wage-employed has a profile $(d_i, y_i) = (1,0)$. According to this profile, the woman participates in the paid labor force ($d_i = 1$), earning an income as wage-employed worker ($y_i = 0$). Combining (3.7) and (3.8) yields the total income of such a household as follows:

$$R(1,0,X_i) = \theta_i M_i + AE_i - \underline{c} \tag{3.20}$$

for all i, such that $(d_i, y_i) = (1,0)$.

From (3.1), substituting in (3.2), (3.3), (3.9), and (3.20) yields the utility payoff of wage employment as follows:

$$U^i(1,0,X_i) = \log\left(\theta_i M_i + AE_i - \underline{c} - p_c K_5^i - c_i^k K_5^i\right) + \gamma\beta\eta \log c_i^k \\ + (1-\eta)\gamma\beta \log(1+\delta_0\ E_i) + \gamma(1-\beta)\log K_5^i \tag{3.21}$$

for all i, such that $(d_i, y_i) = (1,0)$.

This household's problem thus reduces to:

$$\max_{\langle c_i^k \rangle} U^i(1,0,X_i).$$

The interior solution to this problem writes as follows:

$$c_i^k = \frac{\gamma\eta\beta}{(1+\beta\gamma\eta)K_5^i}\left[\theta_i M_i + AE_i - \underline{c} - p_c K_5^i\right] \tag{3.22}$$

$$c_i^a = \frac{1}{1+\beta\gamma\eta}\left[\theta_i M_i + AE_i - \underline{c} - p_c K_5^i\right]. \tag{3.23}$$

From (3.21), substituting in (3.22), and (3.23), rearranging terms yields the utility payoff of wage employment as follows:

$$U^{i*}(1,0,X_i) = (1+\gamma\beta\eta)\log\left[\theta_i M_i + AE_i - \underline{c} - p_c K_5^i\right] \\ + (1-\eta)\gamma\beta\log(1+\delta_0\ E_i) + \gamma\left[1-\beta(1+\gamma\eta)\right]\log K_5^i + Z \tag{3.24}$$

for all i such that $(d_i, y_i) = (1,0)$.

Just as in the case of self-employment, marriage has a positive effect on the payoff of wage employment,

$$U^{i*}(1,0,X_i)\mid_{M_i=1} - U^{i*}(1,0,X_i)\mid_{M_i=0} > 0,$$

as does education, $\partial U^{i*}(1,1,X_i)/\partial E_i > 0$. Furthermore, a close inspection of expression (3.24) reveals that motherhood has an ambiguous effect on this payoff, due to childcare costs.

3.4.5 Optimal Occupational Decisions

As stated above, each household must choose the occupation profile of the woman, $(d_i, y_i) \in \{0,1\} \times \{0,1\}$, to maximize its utility payoff:

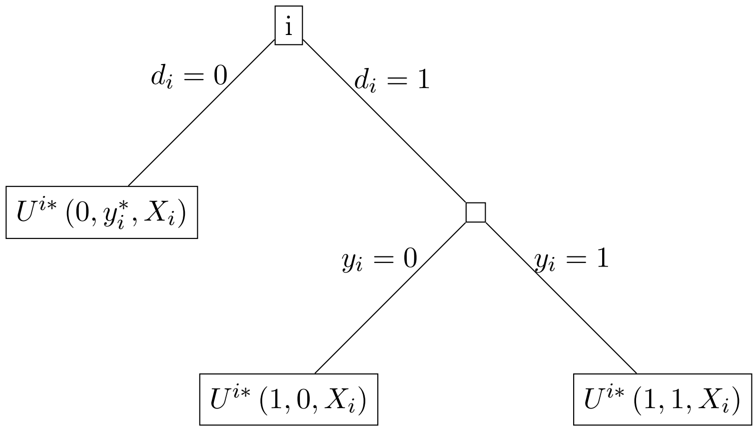
$$\max_{d_i}\left\{\max_{(y_i)} U^{i*}(d_i, y_i, M_i)\right\}$$

where

$$U^{i*}(1,y_i,X_i) = \begin{cases} U^{i*}(1,0,X_i) & \text{if } y_i = 0 \\ \\ U^{i*}(1,1,X_i) & \text{if } y_i = 1 \end{cases}.$$

The decision tree characterizing this two-stage problem is represented below (Figure 3.1), along with the associated utility payoffs:

Figure 3.1: Decision tree



We apply a backward induction process to solve this two-stage problem. First, given that in the first stage the woman plans to be employed, in the second stage the household then must choose the type of employment this woman will perform,y_i: this amounts to solving the following problem:

$$\max_{(y_i)} U^{i*}(1, y_i, M_i).$$

Next, given the optimal decision at this second stage, y_i^*, the household then decides whether or not the woman will participate in the paid labor force. This amounts to solving the following decision problem:

$$\max_{d_i} \hat{U}^{i*}(d_i, y_i^*, M_i),$$

where $y_i^* \equiv \arg\max_{y_i} U^{i*}(1, y_i, M_i)$. We therefore solve the second stage first. We formally apply this backward induction process in what follows.

Optimal Choice of Employment Type

Given that the woman in household i plans to be employed, what type of paid employment will her household select? Since selection of employment type is a binary decision in our model, household i will make this selection by balancing between the utility payoff of self-employment and that of wage employment.

Denote as $Y_i^* := \hat{U}^i(1, 1, X_i) - \hat{U}^i(1, 0, X_i)$, the net payoff gain of selecting self-employment as the female spouse's paid employment. Using (3.19) and (3.21), this net payoff rewrites as follows:

$$Y_i^* = \bar{\beta} \log\left[\frac{\theta_i M_i + \underline{A} E_i - \nu \underline{A} K_5^i E_i}{\theta_i M_i + A E_i - \underline{c} - p_c K_5^i}\right] - \underline{\beta} \log\left(\frac{1 + \delta_0\ E_i}{1 + \delta_1\ E_i}\right) = S(E_i, K_5^i, M_i), \tag{3.25}$$

where $\bar{\beta} := 1 + \gamma\beta\eta$ and $\underline{\beta} = (1-\eta)\,\gamma\beta$.

Households for which female self-employment (i.e., $y_i = 1$) is optimal are those in which the woman has individual characteristics X_i such that

$$Y_i^* > 0. \tag{3.26}$$

Households that choose wage employment (i.e., $y_i = 0$) are those in which the woman has individual characteristics X_i such that

$$Y_i^* < 0. \tag{3.27}$$

(i) Does motherhood influence women's self-employment?

To address this issue, we compare the net payoff of self-employment for a mother of $K_5^i > 0$ children to the net payoff of self-employment for a childless, but otherwise identical, woman. In other words, we are interested in the sign of the difference $\Delta Y_K^* = S(E_i, K_5^i, M_i) - S(E_i, 0, M_i)$. Using (3.25), this difference can be shown to be equivalent to:

$$\Delta Y_K^* = \underline{c}\nu\underline{A}K_5^i E_i + p_c K_5^i(\theta_i M_i + AE_i) - \nu\underline{A}K_5^i E_i(\theta_i M_i + \underline{A}E_i). \tag{3.28}$$

Given the search cost $\underline{c}$ and the cost of nonparental caregiving p_c, the sign of this effect depends on a woman's level of education E_i and her marital status M_i. Clearly, this sign is ambiguous, implying that the effect of motherhood on women's self-employment probabilities is an empirical issue. As a result, one expects the socioeconomic context in which women's self-employment takes place to determine whether or not motherhood matters to women's self-employment. In particular, a close inspection of (3.28) yields the following result:

Proposition 3.4.1 *Let Assumptions 1 and 2 hold simultaneously. If*

$$p_c \geq \nu\underline{A}\bar{E}, \tag{3.29}$$

where $\bar{E} = \max_{(i)} E_i$*, then, motherhood encourages participation in self-employment:* $\Delta Y_K^* > 0$ *.*

Condition (3.29) implies that the motherhood penalty in wage employment exceeds its counterpart in self-employment. Where this condition holds, Proposition 2.1 states that motherhood pulls women into self-employment. In particular, in the African setting where parents can rely on kinship ties to secure affordable nonparental care for their younger children (e.g., the child grand-mother, or another relative), condition (3.29) is likely to be violated. By contrast, in developed countries where the concept of nuclear families is predominant, this condition is more likely to be satisfied in the absence of government subsidization of nonparental caregiving to preschool children.

(ii) Does marriage influence women's self-employment?

To explore the effect of marriage on women's self-employment, we compare the net payoff of self-employment for a married woman ($M_i = 1$) to the corresponding net payoff for a single, but otherwise identical, woman ($M_i = 0$). For this purpose compute the difference $\Delta Y_M^* = S(E_i, K_5^i, 1) - S(E_i, K_5^i, 0)$, using (3.25). We can then show that this difference is equivalent to:

$$\Delta Y_M^* = \theta_i \left[\left(A - \underline{A}\right) E_i - \underline{c} - \left(p_c - \nu \underline{A} E_i\right) K_5^i \right]. \tag{3.30}$$

Under Assumption 2, the sign of (3.30) is clearly ambiguous, as it depends on the woman's other characteristics (including her level of education and the number of children she has), as well as the characteristics of the socioeconomic environment in which she lives (including the search cost for wage employment and the cost of nonparental caregiving).

The above comparative statics implies that the question of whether marriage influences women's self-employment does not have a clear theoretical answer, as so many factors play a role in mediating the effect of marriage.

(iii) Does education influence women's self-employment?

To clarify as much as possible the picture of the effect of education on women's self- employment, we assume education is a continuous variable and take the partial derivative of expression (3.25) with respect to E_i yields:

$$S_E = -\frac{\bar{\beta}\left(\left[A - \left(1 - \nu K_5^i\right) \underline{A}\right] \theta_i M_i + \underline{c} + p_c K_5^i\right)}{\left[\theta_i + \left(1 - \nu K_5^i\right) \underline{A} E_i\right] \left[\theta_i + A E_i - \left(\underline{c} + p_c K_5^i\right)\right]} + \frac{\underline{\beta}\left(\delta_1 - \delta_0\right)}{\left(1 + \delta_1 \; E_i\right)\left(1 + \delta_0 \; E_i\right)} \tag{3.31}$$

Note then that, under Assumptions 1 and 2, the first term of the partial derivative in (3.31) is strictly negative. Since $\delta_1 > \delta_0$, in the presence of a balancing act motive for self-employment as implied by (3.6), clearly, expression (3.31) has an ambiguous sign. This implies that the question of whether education influences women's self-employment has no clear theoretical answer, thus making it an empirical issue. In particular, in socioeconomic environments where

$$\underline{\beta}\left(\delta_1 - \delta_0\right) \to 0, \tag{3.32}$$

the partial derivative in (3.31) is strictly negative: $S_E < 0$. Given that by construction, $\underline{\beta} \in (0, 1)$, condition (3.32) implies that the gain in terms of child quality of switching from wage employment to self-employment (i.e., the balancing act motive for women's self-employment) exists, but has only a modest effect on child quality. This condition is likely to hold in a typical sub-Saharan African society, where nonparental childcare is mainly provided by live-in nannies, often recruited among members of the extended family (including the child paternal or maternal grand-mother, or other female relatives from the extended family network, which can help the child emotionally). By contrast, in a typical developed country where the dominant family model is a nuclear family, it is not clear that a condition such as (3.32) can hold. Hence the following proposition:

Proposition 3.4.2 *Let Assumptions 1 and 2 hold simultaneously, along with conditions (3.6) and (3.32). Then, being educated discourages participation in self-employment.*

Proposition 2.2 is our main result. It implies that in a typical developing economy, education has a negative effect on a woman's probability of self-employment.

Note, however, that the above propositions are conditional on women's decision to participate in the paid labor force. In particular, Proposition 2.2 is a prediction on how working women sort into different types of employment on the basis of their respective levels of education in a developing country's context, but not necessarily a prediction on how women in general behave with respect to this issue, including those who chose not to participate in the labor market. Hence, our focus on the labor force participation decision.

Optimal Participation Decision

What factors promote women's participation in the labor force? Are these factors identical to those that influence the type of employment they engaged in? We address these issues by analyzing the first stage labor force participation decision problem facing each household. This decision problem rewrites as follows:

$$\max\left\{U^{i*}\left(0,y_i^*,X_i\right);\hat{U}^{i*}\left(1,y_i^*,X_i\right)\right\}$$

where $U^{i*}\left(0,y_i^*,X_i\right)$ denotes the utility payoff to non participation defined in (3.13), and

$$\hat{U}^{i*}\left(1,y_i^*,X_i\right)=\max_y U^{i*}\left(1,y_i,X_i\right)$$

is the utility payoff of participation (i.e., $d_i=1$), when the type of employment has been optimally selected by the household in the second stage. Since the labor force participation decision is binary, each household makes this choice by comparing the two payoffs.

Consider the net payoff of participation (i.e., the difference between the payoff of participating and the payoff of non participation). This net payoff, denoted D_i^*, writes as follows:

$$D_i^*\ :=\hat{U}^{i*}\left(1,y_i^*,X_i\right)-U^{i*}\left(0,y_i^*,X_i\right).$$

Since a household headed by a single woman ($M_i=0$) cannot survive without participating in the paid labor force, it follows that for such a household, the net payoff gain of participation is $D_i^*|_{M_i=0}=+\infty$.

By contrast, for co-headed households ($M_i = 1$), we obtain the net payoff of participation as follows using (3.13), (3.19), and (3.24):

$$D_i^*|_{M_i=1} = \begin{cases} \bar{\beta} \log \left[\dfrac{\theta_i + AE_i - \underline{c} - p_c K_5^i}{\theta_i} \right] - \underline{\beta} \log \left(\frac{1+E_i}{1+\delta_0 E_i} \right) & \text{if } y_i^* = 0 \\ \\ \bar{\beta} \log \left[\dfrac{\theta_i + \underline{A}E_i - \nu \underline{A} K_5^i E_i}{\theta_i} \right] - \underline{\beta} \log \left(\frac{1+E_i}{1+\delta_1 E_i} \right) & \text{if } y_i^* = 1 \end{cases} \tag{3.33}$$

for all i such that $M_i = 1$. Since $\theta_i > 0$ for all co-headed households, it is clear that $D_i^*|_{M_i=1}$ is finite. Married women who participate in the paid labor force (i.e., $d_i = 1$) have individual characteristics X_i such that

$$D_i^*|_{M_i=1} > 0. \tag{3.34}$$

Those who do not participate (i.e., $d_i = 0$) have individual characteristics X_i such that

$$D_i^*|_{M_i=1} < 0. \tag{3.35}$$

Recall that expression (3.33) is written for married women. An important issue to consider therefore is whether marriage impedes women's participation in the paid labor force. A close inspection of this expression indicates that its sign depends on a married woman's individual characteristics (level of education and motherhood status) as well as on socioeconomic factors such as the cost of nonparental childcare and of searching for a wage employment. In other words, depending on the number of children a woman, her level of education, or the type of job she would like to hold, the sign of this net payoff of participation may be positive or negative.

To analyze the effect of motherhood and education on women's participation in the paid labor force, it suffices to compare the net payoff of participation for a mother of $K_5^i > 0$ children to the net payoff of participation for a childless, but otherwise identical, woman. Using (3.33), it can be shown that this difference writes as follows:

$$\Delta D_i^*|_{M_i=1} = \begin{cases} \bar{\beta} \log \left[\dfrac{\theta_i + AE_i - \underline{c} - p_c K_5^i}{\theta_i + AE_i - \underline{c}} \right] & \text{if } y_i^* = 0 \\ \\ \bar{\beta} \log \left[\dfrac{\theta_i + \underline{A}E_i - \nu \underline{A} K_5^i E_i}{\theta_i + \underline{A}E_i} \right] & \text{if } y_i^* = 1 \end{cases} \tag{3.36}$$

Clearly, for all y_i^*, $\Delta D_i^*|_{M_i=1} < 0$.

Finally, with to the effect of a woman's level of education on the participation decision, taking the partial derivative of (3.33) with respect to yields

$$\frac{\partial}{\partial E_i} D_i^*|_{M_i=1} = \begin{cases} \dfrac{A}{\theta_i + AE_i - \underline{c} - p_c K_5^i} - \dfrac{\underline{\beta}\,(1-\delta_0)}{(1+E_i)\,(1+\delta_0\ E_i)} & \text{if } y_i^* = 0 \\ \\ \dfrac{\left(1 - \nu K_5^i\right) \underline{A}}{\theta_i + \underline{A}E_i - \nu \underline{A} K_5^i E_i} - \dfrac{\underline{\beta}\,(1-\delta_1)}{(1+E_i)\,(1+\delta_1\ E_i)} & \text{if } y_i^* = 1 \end{cases}$$

As long as parental and nonparental caregiving are imperfect substitutes (i.e., in the sense that $\delta_0 < \delta_1 \leq 1$), the above derivative has an ambiguous sign. This is because female education generates a labor market return that trades off its return in terms of child development. On one hand an educated woman who participates in the labor force bring home a higher income. But on the other hand, when nonparental caregiving imperfectly substitutes for parental caregiving, an educated woman who participates in the paid labor force adversely impact her younger children's cognitive and socioemotional development. Because of these contrasting effects, education has an ambiguous effect on a woman's participation in the paid labor force. The following proposition summarizes the above discussion:

Proposition 3.4.3 *Let Assumptions 1 and 2 hold simultaneously. Then, motherhood has a negative effect on a woman's participation in the paid labor force. However, education and marriage have ambiguous effects on this participation.*

To summarize, our model of how women's individual characteristics affect their employment outcomes leads to one clear prediction only in a developing economy context: motherhood has a negative effect on a woman's probability of participation in the paid labor force. With respect to the issue of which individual characteristics matter to women's self-employment, our model suggests that it has no clear theoretical answers in a developing economy context. Motherhood and education each have two opposite effects on a woman's net payoff of self-employment: a positive and a negative effect. Which of these two is dominant therefore is empirical issue. This motivates our empirical analysis to follow.

3.5 Data

To investigate the effect of women's individual characteristics on their self-employment probabilities, we construct our main sample using data from Uganda's 2016 *Demographic and Health Surveys* (DHS). The DHS database contains information about respondents' socio-economic characteristics such as their education level, employment status, age group, marital status, birth history, household characteristics (e.g., household size, number of children, etc.), as well as their area and region of residence. Our sample contains 15 683 women aged between 15 and 49. We excluded any women with missing information on our key variables (employment status, the number of preschool children and/or educational attainment).

(i) Measuring Labor Force Participation and self-employment

We use the women's questionnaire to construct two main outcome variables: the employment and self-employment indicator functions. Employment is measured using an employment indicator function equal to 1 if a woman reported working in the last 12 months for pay, and 0 otherwise. Self-employment is denoted using an indicator function equal to 1 if a working woman reported being self-employed for the last 12 months, and 0 if she reported being wage-employed. We use the ILO definition of self-employment. The ILO defines self-employed workers as persons of working age who are in one of the following categories: (a) employers; (b) own-account workers (or self-employed without hired employees); (c) contributing family workers and members of producers' cooperatives.[6] A wage-employed woman is defined as someone who reported working for pay in a category other than one of the three self-employed categories.

(ii) Measuring Marital Status, Motherhood, and Education

The questionnaire also contains a number of covariates of interest, including: marital status, motherhood, and educational attainment. The marital status indicator function equals 1 if a woman reported being married or cohabitating with a male partner for the last 12 months, and 0 if she reported being single. Motherhood is measured by the number of preschool children a woman has, while her education level is measured by the number of years of schooling she completed. This way of measuring women's individual characteristics is standard in the literature (Agüero and Marks, 2011; Heath, 2017).

3.5.1 Data and Descriptive Facts

Table (3.1) reports summary statistics in the sample for key variables at both the individual and the regional levels. Wage-employed women have the highest level of education, with 8, 24 years of schooling completed on average. This is more than double the average years of schooling among the self-employed (3.89), or the non-employed (3, 75). The reverse is true for the motherhood variable. The non-employed women have a greater number of preschool children on average (0.73), whereas the wage-employed women have only (0.2). Self-employed women have on average nearly three times as many preschool children (0.56) as their wage-employed counterparts. Wage-employed women are also less likely to be married or cohabitating with a male partner (0.267) than self-employed women (0.536) or the non-employed (0.600).

Table (3.1) also presents the proportion of wage-employed, self-employed, and non-employed women in different age groups. The age distributions of self-employed and wage-employed women are almost

[6] see https://www.ilo.org/global/statistics-and-databases/statistics-overview-and-topics/status-in-employment/current-guidelines/lang–en/index.htm

identical. Importantly, self-employment is predominantly a feature of the rural, mostly informal, economy. Nearly 68 percent of all self-employed women in our analysis sample reside in rural areas, compared to only 32 percent in urban areas. The reverse is true for wage employment, with roughly 83 percent of all wage-employed women residing in urban areas.

3.5.2 Stylized Facts on Women's Employment Outcomes

Women's summary statistics in our data show six worthy stylized facts related to their employment outcomes:

(i) Women have much higher self-employment rates than men

Figure (3.2) shows that women are much more likely than men to be self-employed. Moreover, these high self-employment probabilities persist across all age groups.[7]

(ii) Married women and women with younger children have higher self-employment rates

Figures (3.3) and (3.4) show self-employment by marriage and motherhood status. Figure (3.3) suggests that being married is positively correlated with women self-employment. On average, roughly 84 percent of working women married or cohabitating with a male partner are self-employed, compared to roughly 67 percent of single working women, a difference of roughly 17 percentage points. However, even for single women, at 67 percent, the proportion of self-employed is still well above that of men in our sample (46.66 percent).

Figure (3.4) shows that the proportion or working women who are self-employed is higher for those with more preschool children, even reaching 100 percent for mothers of four or more children. Yet, at 70 percent, the proportion of childless self-employed women is still very large, casting doubt on motherhood as a determining factor of women's self-employment.

[7] This is typical of Sub-Saharan Africa, which exhibits some of the highest labor force participation rates for women in the world. In 2017, 63 percent of women aged 15 or higher were active participants in the labor force, higher than any other region in the world (International Labour Organization, ILOSTAT database, Data retrieved in September 2018). However, the bulk of this participation takes place in low-pay self-employment (Chen, 2001; Heintz and Valodia, 2008; Ulrichs, 2016). For instance, according to a 2016 ILO report, self-employment accounts for around 90 percent of women workers in sub-Saharan countries such as Ethiopia, Benin, Cote d'Ivoire, Mali and Niger. The self-employment probabilities of women in our analysis sample echoes this trend, with 77.1 percent of working women reporting being self-employed over the last twelve months, compared to only 46.6 percent of men, a difference of 30.4 percentage points. Given that self-employment in developing countries takes place predominantly in the low-pay informal sector, and self-employed women tend to earn much less than self-employed men (Bank, 2011), understanding why women have such high self-employment probabilities in developing countries is key to understanding why the gender gap in pay persists in these countries.

(iii) Women's self-employment decreases with education

Figure (3.5) plots the proportion of self-employed women by age and by levels of education. For all age groups, the proportion of women who are self-employed decreases as education levels increase.

Among working women with no education, the proportion of those who are self-employed is greater than 90 percent. For working women with a high school education, the corresponding figure drops to roughly 62 percent, for working women with more than a high school education, it drops even further to 33 percent. These stylized facts raise the important question of whether lower levels education among Ugandan working women drive their self-employment, and thus motivate our formal investigation of the causal effect of education on women's self-employment probabilities.

(iv) Women have much lower employment probabilities than men

If all women in our analysis sample where employed, a relevant econometric model for obtaining consistent estimates of the causal effects of motherhood would be one that corrects for endogeneity. However, Figure (3.6) shows a large gender gap in labour force participation. Men are much more likely to participate than women. Only 33.3 percent of women in our sample held a job for the last year, compared to 89.3 percent of men, for the same period, a gender gap of more than 56 percentage points. Moreover, this gap is present and relatively stable at all age groups in our sample. This low female labor force participation rate in our analysis sample raises the issue of whether working women's self-selection into employment is nonrandom, and motivates our formal investigation of this issue as an integral part of our empirical strategy.

3.6 Empirical Strategy

In this section we outline our empirical strategy for testing the main prediction of our theoretical model that education is a driver of women's self-employment in a developing economy. In particular, the theory developed in the previous section suggests that a woman becomes self-employed if her household's net utility payoff from picking this employment option is strictly positive. A linear approximation of the net utility payoff of self-employment can be written as function of observed and unobserved factors, as follows:

$$y_{itr} = \mathbb{1}\left[\beta_0 + \gamma_t + \gamma_r + \beta_1 E_{itr} + \beta_2 K_5^{itr} + \beta_3 M_{itr} + \beta_4 R_{itr} + \epsilon_{itr} > 0\right], \tag{3.37}$$

where $y_{itr} \in \{0, 1\}$ denotes the self-employment status of a working woman i from birth cohort t residing in region r, and $\mathbb{1}[\cdot]$ is an indicator function equal to 1 if the inequality in parentheses holds true for woman i from birth cohort t residing in region r and 0 otherwise; E_{itr} is her level of education measured in years of schooling completed; K_5^{itr} is the number of preschool children she has; M_{itr} her marital status, and R_{itr} the vector of other individual and household level controls (including ethnic group and area residence). The term ϵ_{itr} is a mean-zero error term which captures the effect of unobserved factors that influence women's employment outcomes; γ_t and γ_r are respectively the birth cohort and the region fixed effects

representing changes in self-employment risks related to her birth cohort t and to self-employment opportunities in her region of residence r.

If all the covariates in the regression described in Equation (3.37) were exogenous, and all women in our empirical setting were employed, estimates obtained from (3.37) would be unbiased, thus yielding a correct identification of the causal of effect of motherhood and education. However, the estimation of the self-employment equation in (3.37) raises two important challenges. First, in our analysis sample, labor force participation among working-age women is below 50 percent, which raises the issue of whether selection into the labor force is nonrandom. Second, there is the potential endogeneity of our main covariates of interest namely, education and motherhood.

3.6.1 Sample Selection Bias

Equation (3.37) is specified for the entire active population of women. However, the self- employment outcome is only observed among working women. Indeed, in our sample only 33.3 percent of working-age women are employed. Table (3.2) highlights differences in individuals characteristics between the sub-sample of working women and the sub-sample of those who are not gainfully employed. These differences do support the presence of sample selection bias in our empirical setting.

In our theory section a selection rule is formulated by the net utility payoff of labor force participation whereby a woman participates in the paid labor force if and only if this net utility payoff is strictly positive. In this context, whether or not the self-employment outcome, y_{itr}, is observed depends on the following employment outcome:

$$d_{itr} = \mathbb{1}\left[\alpha_0 + \lambda_t + \lambda_r + \alpha_1 E_{itr} + \alpha_2 K_5^{ijr} + \alpha_3 M_{itr} + \alpha_4 X_{itr} + \eta_{itr} > 0\right] \tag{3.38}$$

where $\mathbb{1}[\cdot]$ is an indicator function equal to 1 if, for woman i of birth cohort t residing in region r, the (linearized) net utility payoff of participation in the paid labor force (i.e., the term in parentheses) is strictly positive, and zero otherwise;

η_{itr} is a zero-mean error term which captures the effect of unobserved factors that influence a woman's labor force participation decision [8]; λ_t and λ_r are respectively the birth cohort and the region's fixed effects for labor force participation, representing the changes in woman i's labor force participation related to her birth cohort t and to employment opportunities in her region of residence r.

A sample selection bias thus occurs when there exists a statistically significant correlation between unobservables in the errors terms in Equations (3.37) and (3.38). If sample selection were the only threat to the identification of the causal effects of motherhood and education on a woman's self-employment probability, estimation of the two-equation model represented by Equations (3.37) and (3.38) would produce consistent estimates of these effects.

3.6.2 The Endogeneity of Education and Motherhood

As argued in Semykina (2018), the literature on women's employment outcomes finds evidence of joint determination between fertility and women's labor force participation and labor supply, implying that a similar problem is likely to arise in the context of women's self-employment decisions. Likewise, the labor literature finds evidence of joint determination between an individual's education and employment decisions (Ahituv and Tienda, 2004), implying that this problem is also likely to arise in the context of women's self-employment decisions. We describe our strategy for addressing these two types of identification challenges.

A. Eligibility to Universal Primary Education

We take advantage of a natural experiment created by the Uganda's *Universal Primary Education* (UPE) reform to obtain exogenous variation in the level of education. Exploiting school reforms to create instruments for education levels has proven a popular empirical strategy to investigate the causal effect of education on outcomes related to health (Brunello et al., 2016), fertility (Monstad et al., 2008; Osili and Long, 2008; Cygan-Rehm and Maeder, 2013; Geruso et al., 2011; Fort et al., 2016) and employment (Denny and Harmon, 2000; Duflo, 2001).

In 1997, the Ugandan government implemented the UPE reform to remove financial barriers to primary school (Deininger, 2003; Grogan, 2008). This policy abolished both direct and indirect school fees (such as donations to parents funds and school uniforms) which impeded school access for the poor. In order to promote the program, information dissemination campaigns were launched across the country. These campaigns also included target messages that specifically focused on educating girls in rural areas. Available evidence points to the success of these campaigns. Deininger (2003) and Grogan (2008) find that the elimination of school fees under the UPE policy greatly increased the

[8] Heckman (1976, 1979) sample selection model assumes that the error terms ϵ_{itk} and η_{itr} are jointly normally distributed.

probability that pre-school age children enrolled in school in Uganda. However, the pre vs. post-reform specification of the instrumental variable for education is incomplete as it does not account for eligible children who were already enrolled in primary school when the reform was introduced. Upon the launching of the UPE reform, a maximum of seven years of primary school were available to every preschool child. Thus, children aged 6 or less in 1997 had full eligibility. For children aged 7 or higher in 1997, however, school fees were eliminated only for the rest of their primary education. Children who had already completed their primary school at that time therefore were not eligible, and thus could not benefit from the fees' removal.

Our instrumental variable for education is a measure of the eligibility of each school-age child i from birth cohort t to the UPE reform. We measure this eligibility as the number of years an individual can benefit from the UPE reform, given his date of birth. Our instrument for education thus takes the form of an eligibility function, denoted $EligibilityUPE$, whose main argument is the individual's year of birth:

$$EligibilityUPE_{it} = \min\left[\max\left\{\left(SA + \bar{N}\right) - \left(N_1 - N_{it}\right), 0\right\}, \bar{N}\right], \tag{3.39}$$

where SA represents the primary school enrolment age in Uganda (which is age six); $\bar{N}$, the maximum number of years an individual can benefit from the UPE (which is seven years of primary education), N_{it}, the individual's year of birth, and N_1, the date of the implementation of the UPE (which is 1997). Note, $EligibilityUPE_{ij} \in \{0, 1, 2, 3, 4, 5, 6, 7\}$. An individual's level of exposure to the UPE reform thus is clearly exogenous and random.

Given this definition of the instrumental variable $EligibilityUPE_{ij}$ (see Equation 3.39) for a woman's level of education, E_{ij}, we can specify the reduced form equation to estimate the effect of $EligibilityUPE_{ij}$ on E_{ijr} as follows:

$$E_{ijr} = a_0 + \theta_j + \theta_r + a_1 EligibilityUPE_{ij} + a_2 X_{ijr} + \upsilon_{ijr} \tag{3.40}$$

where a_1 captures the effect of the UPE program on the total number of years of schooling for women who benefited from this program. Equation (3.40) allows us to estimate the exogenous increase in schooling generated by the UPE reform. θ_t and θ_r are respectively the birth cohort and regional fixed effects that may influence women's educational attainment.

B. Exposure to Infertility shocks

Following Agüero and Marks (2011), we use the union of two infertility measures derived from survey respondents' self-reported infertility. The first measure uses reported infertility as the reason for not currently using contraceptives (Infertile1). The second measure uses reported inability to have children as the reason for not achieving their desired number of children (Infertile 2). Our infertility indicator is:

$$Infertility = \max\{Infertility1, Infertility2\}.$$

Unlike instruments for fertility which target only women with children (such as twins at first birth (Bronars and Grogger, 1994) or the sex composition of the first two children (Angrist and Evans, 1998; Cruces and Galiani, 2007)) this instrument enables both women with, and without, children to be included in the sample. We do restrict our sample to women who are sexually active and reported that they do not use contraceptives. As a result, the size of our sample drops from 15497 to 12371 women, a lost of only roughly 20.17 percent. In this sub-sample, we observe fewer preschool children in the families of infertile women (0.17), than for their fertile counterparts (0.69).[9]

We estimate the following reduced form equation to measure the effect of infertility shocks on the number of preschool children a woman has:

$$K_5^{ijr} = b_0 Infertility_{ij} + b_1(1 - Infertility_{ij}) + \psi_j + \psi_r + b_2 T_{ijr} + \nu_{ijr} \tag{3.41}$$

where K_5^{ijr} is the number of preschool children of woman i of birth cohort j residing in region r, $Infertility_{ij}$ is her infertility status, ψ_j and ψ_r represent the fixed effects of the birth cohort and the region on women's fertility. T_{ijr} include control variables such as the ethnic group, the family wealth index and the area of residence. The parameter β_0 captures the effect of being exposed to infertility on the number of preschool children. According to Agüero and Marks (2011), this effect is expected to be negative and statistically significant.

3.6.3 Estimation Method

To obtain consistent estimators of vectors α, β, $\mathbf{a}$, and $\mathbf{b}$, we estimate the following four-equation econometric model:

$$y_{itr} = \mathbb{1}\left[\beta_0 + \gamma_t + \gamma_r + \beta_1 E_{itr} + \beta_2 K_5^{itr} + \beta_3 M_{itr} + \beta_4 R_{itr} + \epsilon_{itr} > 0\right] \tag{3.42}$$

$$d_{itr} = \mathbb{1}\left[\alpha_0 + \lambda_t + \lambda_r + \alpha_1 E_{itr} + \alpha_2 K_5^{itr} + \alpha_3 M_{itr} + \alpha_4 X_{itr} + \eta_{itr} > 0\right] \tag{3.43}$$

$$E_{itr} = a_0 + \theta_k + \theta_t + a_1 EligibilityUPE_{it} + a_2 X_{itr} + \upsilon_{itr} \tag{3.44}$$

$$K_5^{itr} = b_0 Infertility_{it} + b_1(1 - Infertility_{it}) + \psi_t + \psi_r + b_2 T_{itr} + \nu_{itr} \tag{3.45}$$

[9] Aguero and Marks (2008), Field and Ambrus (2008), Maccini and Yang (2009), Agüero and Marks (2011), and Shah and Steinberg (2017) argue that measurement errors do not present a significant problem for self-reported fertility shocks. An additional concern is that the presence of male family members at the time women answered the survey questionnaire may lead these women to misreport their fertility status. Our empirical results showed no evidence that the presence of other family members was correlated with responses to fertility questions (Results of this test are reported in Table (3.3)); see also Agüero and Marks (2011).

(3.42) is the self-employment (or main) equation. It describes the self-employment status, y_{itr}, of woman i of birth cohort t residing in region r, which is equal to 1 if self-employed and zero otherwise. Equation (3.43) is the employment (or selection) equation, describing women's selection into paid employment, such that $d_{itr} = 1$ if she participates in the paid labor force and zero otherwise.

Equations (3.44) and (3.45) are first stage equations that account for education and motherhood being potentially endogenous regressors. Estimation of this four-equation model is based upon the assumption that there is a correlation between errors terms in Equations (3.42) and (3.43), between error terms in Equation (3.42) and in each of the first-stage equations, as well as between error terms in Equation (3.43) and in each of the first-stage equations. Were it otherwise, consistent estimators could have been obtained by estimating each equation separately.

To implement the estimation of this four-equation model, we combine the Full Information Maximum Likelihood Estimation method with a control function approach for endogenous regressors. Our estimation of this four-equation model unfolds in two steps. In a first step, we estimate the reduced form (First stage) equations (3.40) and (3.41). The estimated residuals are then inserted in the second stage regressions as additional regressors, as in Terza et al. (2008). However, since both education and the number of preschool children (our outcomes of interest in first-stage regressions) are count variables, the normality requirement for the use of the control function approach may not be satisfied. To address this issue, we compute *Anscombe residuals* from the fitted Poisson regressions, using the Anscombe transform that approximate a non-normally distributed variable with one that follows a Gaussian distribution (Anscombe, 1948). The Anscombe residuals thus computed are then inserted as additional covariates in the second-stage regressions (in both the selection and employment type equations). In a second step, we estimate these two equations jointly, using the Full Information Maximum Likelihood Estimator's method.

3.7 Empirical Results

In this section, we report and discuss the results of our estimation of the effects of women's individual characteristics on their self-employment probabilities. To assess the importance of correcting for nonrandom self-selection into the labor force and the endogeneity of both education and motherhood, we follow a three-step approach. First, we estimate the self-employment equation separately (single-equation model), then we estimate a pair of two-equation models of women's self-employment in which we correct for nonrandom self-selection and the endogeneity of motherhood separately. Next, we estimate a three-equation model that corrects for these two problems jointly, as in Semykina (2018). Finally, we estimate the full four-equation model described above, and in which we correct not only for sample selection and the endogeneity of motherhood, but also the endogeneity of education in both main and selection equation.

When we do not correct for any potential identification problem (column 1, Table (3.4)), we find that motherhood has no statistically significant effect on women's self-employment. This result is consistent with our stylized facts suggesting that women's self-employment probabilities are very high (well above 50 percent of working women) irrespective of women's individual characteristics (see Figure (3.4)). However, it is in contrast to Connelly (1992) and Semykina (2018) who find that motherhood has a positive effect on working women's self-employment in a developed economy. We also find that education has a negative effect on a woman's self-employment probability, and this effect is significant. Semykina (2018) finds no effect of education in a developed country's context. However, both results derived from the estimation of Equation 3.37 are only valid if selection into the working women sample is random, and education and motherhood are exogenous.

When we correct only for nonrandom self-selection into the working women sample (see columns 2 and 3, Table (3.4)), results still show that motherhood has no statistically significant effect on women's self-employment. Education also still has a negative and statistically significant effect on self-employment (column 2), albeit with a slightly smaller magnitude compared to the single-equation model (column 1). Furthermore, the correlation coefficient between the error terms in main and selection equations is 0.732 and highly statistically significant (column 3, Table (3.4)). This indicates that sample selection bias exists. However, the results of this estimation are only valid if motherhood and education are both exogenous.

When we correct only for the endogeneity of motherhood (column 4, Table (3.4)), first-stage results show that infertility is negatively correlated with the number of children a woman has, and this effect is highly significant. In particular, the computed F-statistic is large and statistically significant ($F-statistic = 125,31$), implying that there is no evidence of weak instrument. Second-stage results show that motherhood now has a positive and statistically significant effect on women's self-employment: the presence of an additional child increases a woman's self-employment probability by 12 percentage points, which is quite large. However, this result is obtained at the cost of ignoring the presence of sample selection bias reported above.

Next, Table (3.5) report the results of the estimation of a three-equation model in which we account for both nonrandom self-selection into the paid labor force and the endogeneity of the motherhood variable jointly. For this three-equation model, it is assumed that the education variable is exogenous, as in Semykina (2018). We find that motherhood has no statistically significant effect on women's self-employment (column 2, Table (3.5)). This indicates that not correcting for selection and the endogeneity of fertility leads to an overestimation of the effect of motherhood (column 4, Table (3.4)). Correcting for sample selection bias and the endogeneity of motherhood jointly shows that motherhood and marriage only impact a woman's labor force participation probability. In fact, they both reduce this probability (column 3, Table (3.5)). Interestingly, education still has a negative, albeit very small effect on a woman's self-employment (column 2, Table (3.5)). It also has a modest positive effect on a woman's probability of participation in the paid labor force. Both effects are statistically significant and similar in magnitude to those obtained from the estimation of a two-equation model that corrects for nonrandom self-selection only (column 2 and 3 Table (3.4)). These results, however, are only valid if education is exogenous in both the self-employment and the employment outcome equations. We argued above that this is not the case

3.7.1 Estimates for The Full Four-Equation Model

Here, we report and discuss the results of the estimation of our four-equation model that accounts simultaneously for nonrandom self-selection, the endogeneity motherhood, as well as the endogeneity of education in both main and selection equation. Results of this estimation approach are reported in Table (3.6). First, Columns (1) and (2) of Table (3.6) provide strong evidence of the relevance of both our chosen instruments for education and motherhood respectively.[10] We find a positive and statistically significant effect of the UPE reform. The F-statistics is larger than the value 10 recommended by Stock et al. (2002), implying that there is no evidence of weak instrument in the education equation.

For this four-equation model, yet again we find that motherhood has no statistically significant effect on women's self-employment (column 3, Table (3.6)). It only affects women's labor force participation (see column 4 Table (3.6)). Indeed, having an additional preschool child reduces the probability of labor force participation by 7.86 percentage points, which is virtually identical in magnitude to the corresponding value obtained from the estimation of the three-equation model (7.90 percentage points), and in which we do not correct for the endogeneity of the education variable (column 4, Table (3.5)).

Importantly, when compared to the three-equation model of Table (3.5), where we do not correct for the endogeneity of the education variable, education in the full four-equation model not only maintains its negative effect on women's self-employment probabilities, but also, the magnitude of this effect more than triples to reach 3.05 percentage points (column 3, Table (3.6)). Moreover, education no longer has a statistically significant effect on women's labor force participation probabilities (column 4, Table (3.6)), contrary to what we find in column 3, Table (3.5). These contrasting results highlight the importance of correcting for the endogeneity not only of the motherhood variable (Semykina, 2018), but also of the education variable, in both main and selection equations. Not correcting for the latter thus yields inconsistent estimates.

[10]Given that both education and the number of preschool children are count variables, we use *Poisson* regressions to estimate their respective equations.

Results for the estimation of the four-equation model also confirm our findings reported in Table (3.5) that a woman's marital status has no effect on her self-employment probability (column 3, Table (3.6)). Like motherhood, marital status only affects a woman's probability of participation in the paid labor force (column 4, Table (3.6). Indeed, being married reduces a woman's probability of working by roughly 10 percentage points, which is similar in magnitude to the results reported in Tables (3.4) and (3.5).

Compared to Connelly (1992) and Semykina (2018) who find that motherhood drives women's self-employment in developed countries, our results suggest that this effect is not universal, as it does not hold in a developing country context, where self-employment is predominantly a phenomenon of the informal economy characterized by insecure property rights. In this context, women's self-employment is likely to be involuntary, being drive by push factors such as the inability to secure a wage job in the formal sector, due to lack of education.

3.7.2 Robustness Checks

(i) Definition of Motherhood Status

In our main specification, the motherhood variable is a count variable measuring the number of preschool children a working woman has. To assess whether our results are due to this assumption, we adopt an alternative definition of the motherhood variable as an indicator function which equals 1 if a woman has at least one preschool child and zero otherwise. The estimated coefficients of the four-equation model are reported in Table (3.7). Overall, our results still hold. Motherhood has no statistically significant effect on a woman's self-employment probability, while education still has a negative and statistically significant effect on this probability. Both results differ from Semykina (2018) whose study focuses on developed countries.

(ii) Spatial Heterogeneity in The Impact of Motherhood and Education

Our summary statistics indicate that self-employment is more common in rural areas, while wage employment is relatively more common in urban area. We test whether these spatial disparities in women's employment outcomes are a source of heterogeneity in the impacts of women's individual characteristics their self-employment probabilities. We divide our analysis sample into two sub-samples: the rural sub-sample and the urban sub-sample. We estimate the effects of our covariates of interest for each sub-sample. The estimated coefficients are reported in Table (3.8). Test results reveal that education has a negative and statistically significant effect on women's self-employment probabilities in both rural and urban sub-samples. Motherhood, by contrast, has a negative effect on women's labor force participation probabilities only in rural areas. The latter result can be explained by the fact that gendered norms that prescribe homemaking and care-giving as the primary responsibilities of women tend to be stronger in rural than in urban areas.

3.8 Conclusion

We provide new quantitative evidence on the importance of motherhood and education to women's self-employment in a developing economy's context. To guide our empirical analysis, we develop a theoretical model explaining how a woman's individual characteristics affect her labor force participation and self-employment decisions. Our model highlights a key distinguishing feature of self-employment in a developing economy's context: balancing paid work and unpaid child care labor in self-employment may enhance child quality compared to wage employment, when caregiving to younger children is primarily a mother's responsibility. But it also entails a motherhood penalty in the form of loss of labor productivity. This loss of productivity has two different sources. First, there is the fact that self-employment takes place in the informal economy characterized by insecure property rights, which discourages the adoption of productivity-enhancing technologies, thus leading to loss of labor productivity relative to wage employment—a formal sector activity. Second, in self-employment women work while caring for their younger children. Yet caring for younger children while working also reduces a mother's labor productivity. Since wage employment, which takes place predominantly in the formal sector, also entails a motherhood penalty in the form of childcare expenses, having younger children therefore has an ambiguous effect on women's self-employment decisions. This implies that whether children matter to women's self-employment probabilities or not is an empirical issue. Our main contribution therefore is to investigate this empirical issue in the context of a developing economy, using micro-level data from Uganda. Our empirical strategy accounts for the problems of nonrandom self-selection into the paid labor force and the endogeneity of multiple regressors jointly. We obtain consistent estimates of the effects of covariates of interests by estimating a four-equation econometric model, using a Full Information Maximum Likelihood Estimator's method combined with control functions for endogenous regressors. Estimation results separate our study from the existing literature on women's self-employment. First, we find that motherhood has no significant effect on a woman's probability of self-employment in a developing economy's context, which is in contrast to findings that this effect exists and is positive in developed countries (Connelly, 1992; Semykina, 2018). Second, we also find that education has a negative and statistically significant effect on this probability, which is also in contrast to the existing literature showing that education has no statistically significant effect on this probability in a developed country's context (Semykina, 2018).

Overall, our findings suggest that the sociocultural and institutional contexts in which women's self-employment takes place determine the nature of its underlying social factors. In a developing economy, participation in self-employment is more likely to be involuntary than voluntary for women, because it takes place in the informal sector characterized by insecure property rights. In such a context, women's inability to secure wage jobs in the formal sector due to lack of education is the factor that pushes them into low-pay self-employment. To reduce the gender gap in pay in this context, public policy is likely to be more effective if it focuses on promoting female education as well as households' access to quality professional childcare services. Education would draw women out of the informal sector and into the formal sector where the bulk of wage employment takes place, while affordable childcare services would reduce the cost of formal sector employment for mothers.

Bibliography

Agüero, J., M. Marks, and N. Raykar (2012): "The wage penalty for motherhood in developing countries," Tech. rep., Working Paper. Riverside, California: University of California, Riverside.

Aguero, J. M. and M. S. Marks (2008): "Motherhood and female labor force participation: evidence from infertility shocks," *American Economic Review*, 98, 500–504.

Agüero, J. M. and M. S. Marks (2011): "Motherhood and female labor supply in the developing world evidence from infertility shocks," *Journal of Human Resources*, 46, 800–826.

Ahituv, A. and M. Tienda (2004): "Employment, motherhood, and school continuation decisions of young White, Black, and Hispanic women," *Journal of Labor Economics*, 22, 115–158.

Angrist, J. D. and W. N. Evans (1998): "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *The American Economic Review*, 88, 450–477.

Anscombe, F. J. (1948): "The transformation of Poisson, binomial and negative-binomial data," *Biometrika*, 35, 246–254.

Baker, M., J. Gruber, and K. Milligan (2008): "Universal child care, maternal labor supply, and family well-being," *Journal of political Economy*, 116, 709–745.

Baker, M. and K. Milligan (2015): "Maternity leave and children's cognitive and behavioral development," *Journal of Population Economics*, 28, 373–391.

Bank, W. (2011): *World development report 2012: gender equality and development*, World Bank Publications.

Becker, G. S. and H. G. Lewis (1973): "On the Interaction between the Quantity and Quality of Children," *Journal of political Economy*, 81, S279–S288.

Becker, G. S. and N. Tomes (1976): "Child endowments and the quantity and quality of children," *Journal of political Economy*, 84, S143–S162.

Bertrand, M., P. Cortés, C. Olivetti, and J. Pan (2016): "Social norms, labor market opportunities, and the marriage gap for skilled women," Tech. rep., National Bureau of Economic Research.

Bianchi, S. M. (2000): "Maternal employment and time with children: Dramatic change or surprising continuity?" *Demography*, 37, 401–414.

Boserup, E., S. F. Tan, and C. Toulmin (2013): *Woman's role in economic development*, Routledge.

Bronars, S. G. and J. Grogger (1994): "The economic consequences of unwed motherhood: Using twin births as a natural experiment," *The American Economic Review*, 1141–1156.

Brown, A. (2015): "Claiming the streets: property rights and legal empowerment in the urban informal economy," *World Development*, 76, 238–248.

Brunello, G., M. Fort, N. Schneeweis, and R. Winter-Ebmer (2016): "The causal effect of education on health: What is the role of health behaviors?" *Health economics*, 25, 314–336.

Budlender, D. (2008): *The statistical evidence on care and non-care work across six countries*, United Nations Research Institute for Social Development Geneva.

Carneiro, P., K. V. Løken, and K. G. Salvanes (2015): "A flying start? Maternity leave benefits and long-run outcomes of children," *Journal of Political Economy*, 123, 365–412.

Chen, M. A. (2001): "Women and informality: A global picture, the global movement," *Sais Review*, 21, 71–82.

Connelly, R. (1992): "Self-employment and providing child care," *Demography*, 29, 17–29.

Craig, L., A. Powell, and N. Cortis (2012): "Self-employment, work-family time and the gender division of labour," *Work, employment and society*, 26, 716–734.

Cruces, G. and S. Galiani (2007): "Fertility and female labor supply in Latin America: New causal evidence," *Labour Economics*, 14, 565–573.

Currie, J. and E. Moretti (2003): "Mother's education and the intergenerational transmission of human capital: Evidence from college openings," *The Quarterly Journal of Economics*, 118, 1495–1532.

Cygan-Rehm, K. and M. Maeder (2013): "The effect of education on fertility: Evidence from a compulsory schooling reform," *Labour Economics*, 25, 35–48.

Danzer, N. and V. Lavy (2018): "Paid parental leave and children's schooling outcomes," *The Economic Journal*, 128, 81–117.

De La Croix, D. and M. Doepke (2003): "Inequality and growth: why differential fertility matters," *American Economic Review*, 93, 1091–1113.

De la Croix, D. and M. Doepke (2004): "Public versus private education when differential fertility matters," *Journal of Development Economics*, 73, 607–629.

Deininger, K. (2003): "Does cost of schooling affect enrollment by the poor? Universal primary education in Uganda," *Economics of Education review*, 22, 291–305.

Del Boca, D., C. Flinn, and M. Wiswall (2013): "Household choices and child development," *Review of Economic Studies*, 81, 137–185.

Denny, K. and C. Harmon (2000): "Education policy reform and the return to schooling from instrumental variables," *CEPR Discussion Papers 2518, C.E.P.R. Discussion Papers.*

Doss, C. R. (2018): "Women and agricultural productivity: Reframing the Issues," *Development Policy Review*, 36, 35–50.

Drange, N. and T. Havnes (2014): "Early and bright? Child care for toddlers and early cognitive skills," in *Manuscript presented at the EALE Conference in Ljubljana.*

——— (2019): "Early Childcare and Cognitive Development: Evidence from an Assignment Lottery," *Journal of Labor Economics*, 37, 581–620.

Duflo, E. (2001): "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment," *American economic review*, 91, 795–813.

Dustmann, C. and U. Schönberg (2012): "Expansions in maternity leave coverage and children's long-term outcomes," *American Economic Journal: Applied Economics*, 4, 190–224.

Fernández, R., A. Fogli, and C. Olivetti (2004): "Mothers and sons: Preference formation and female labor force dynamics," *The Quarterly Journal of Economics*, 119, 1249–1299.

Field, E. and A. Ambrus (2008): "Early marriage, age of menarche, and female schooling attainment in Bangladesh," *Journal of political Economy*, 116, 881–930.

Fort, M., A. Ichino, and G. Zanella (2017): "The cognitive cost of daycare 0-2 for children in advantaged families," in *Technical Report*, Working Paper.

Fort, M., N. Schneeweis, and R. Winter-Ebmer (2016): "Is education always reducing fertility? Evidence from compulsory schooling reforms," *The Economic Journal*, 126, 1823–1855.

Fu, X., P. Mohnen, and G. Zanello (2018): "Innovation and productivity in formal and informal firms in Ghana," *Technological Forecasting and Social Change*, 131, 315–325.

Garcia-Mainar, I. and V. M. Montuenga-Gomez (2005): "Education returns of wage earners and self-employed workers: Portugal vs. Spain," *Economics of Education Review*, 24, 161–170.

Geruso, M., D. Clark, and H. Royer (2011): "The impact of education on fertility: Quasiexperimental evidence from the UK," Tech. rep., mimeo, Princeton, Princeton University.

Grogan, L. (2008): "Universal primary education and school entry in Uganda," *Journal of African Economies*, 18, 183–211.

Gupta, N. D. and M. Simonsen (2010): "Non-cognitive child outcomes and universal high quality child care," *Journal of Public Economics*, 94, 30–43.

Heath, R. (2017): "Fertility at work: Children and women's labor market outcomes in urban Ghana," *Journal of Development Economics*, 126, 190–214.

Heckman, J. J. (1976): "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models," in *Annals of Economic and Social Measurement, Volume 5, number 4*, NBER, 475–492.

——— (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–161.

Heckman, J. J. and S. Mosso (2014): "The economics of human development and social mobility," *Annu. Rev. Econ.*, 6, 689–733.

Hein, C. and N. Cassirer (2010): *Workplace solutions for childcare*, International Labour Office Geneva, Switzerland.

Heintz, J. and I. Valodia (2008): "Informality in Africa: A review," *Background paper for the Swedish International Development Cooperation Agency (Sida), Unpublished Working Paper.*

Herbst, C. M. (2013): "The impact of non-parental child care on child development: Evidence from the summer participation "dip"," *Journal of Public Economics*, 105, 86–105.

Lim, K. (2018): "Do American mothers use self-employment as a flexible work alternative?" *Review of Economics of the Household*, 1–38.

Maccini, S. and D. Yang (2009): "Under the weather: Health, schooling, and economic consequences of early-life rainfall," *American Economic Review*, 99, 1006–26.

Mercer, J. (2006): *Understanding attachment: Parenting, child care, and emotional development*, Greenwood Publishing Group.

Monstad, K., C. Propper, and K. G. Salvanes (2008): "Education and fertility: Evidence from a natural experiment," *Scandinavian Journal of Economics*, 110, 827–852.

Moock, P. R., P. Musgrove, and M. Stelcner (1989): *Education and earnings in Peru's informal nonfarm family enterprises*, vol. 236, World Bank Publications.

Mousié, R. (2016): "Child care from the perspective of women in the informal economy," *Policy Brief for the UN Secretary-General's High-Level Panel on Women's Economic Empower- ment.*

Noboa-Hidalgo, G. E. and S. S. Urzua (2012): "The effects of participation in public child care centers: Evidence from Chile," *Journal of Human Capital*, 6, 1–34.

Osili, U. O. and B. T. Long (2008): "Does female schooling reduce fertility? Evidence from Nigeria," *Journal of development Economics*, 87, 57–75.

Quisumbing, A. R., K. Hallman, and M. T. Ruel (2007): "Maquiladoras and Market mamas: Women's work and childcare in Guatemala City and Accra," *The Journal of Development Studies*, 43, 420–455.

Rasmussen, A. W. (2010): "Increasing the length of parents' birth-related leave: The effect on children's long-term educational outcomes," *Labour Economics*, 17, 91–100.

Roche, K. (2014): "Female self-employment in the United States: an update to 2012," *Monthly Lab. Rev.*, 137, 1.

Samman, E., E. Presler-Marshall, N. Jones, T. Bhatkal, C. Melamed, M. Stavropoulou, and J. Wallace (2016): "Women's work: Mothers, children and the global childcare crisis," *London: Overseas Development Institute.*

Semykina, A. (2018): "Self-employment among women: Do children matter more than we previously thought?" *Journal of Applied Econometrics*, 33, 416–434.

Shah, M. and B. M. Steinberg (2017): "Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital," *Journal of Political Economy*, 125, 527–561.

Stock, J. H., J. H. Wright, and M. Yogo (2002): "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business & Economic Statistics*, 20, 518–529.

Tertilt, M. (2006): "Polygyny, women's rights, and development," *Journal of the European Economic Association*, 4, 523–530.

Terza, J. V., A. Basu, and P. J. Rathouz (2008): "Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling," *Journal of health economics*, 27, 531–543.

Ulrichs, M. (2016): "Informality, women and social protection: identifying barriers to provide effective coverage," *Overseas Development Institute, Working Paper,*, 435.

Wellington, A. J. (2006): "Self-employment: the new solution for balancing family and career?" *Labour Economics*, 13, 357–386.

Wodon, Q. and C. M. Blackden (2006): *Gender, time use, and poverty in sub-Saharan Africa.*

Appendix

Figure 3.2: Self-employment, by age and by gender

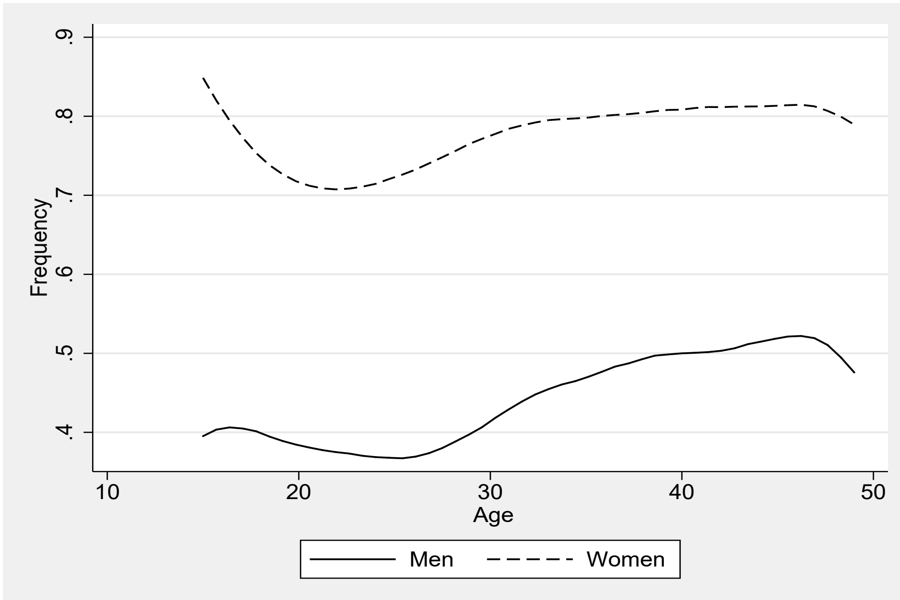


Figure 3.3: Women's self-employment, by age and marital status

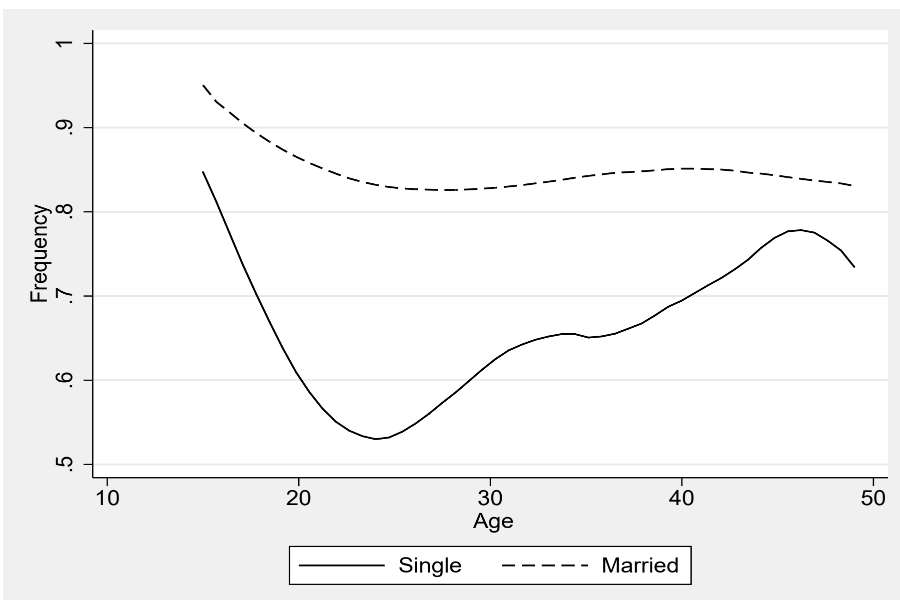


Figure 3.4: Women's self-employment, by age and numbers of preschool children

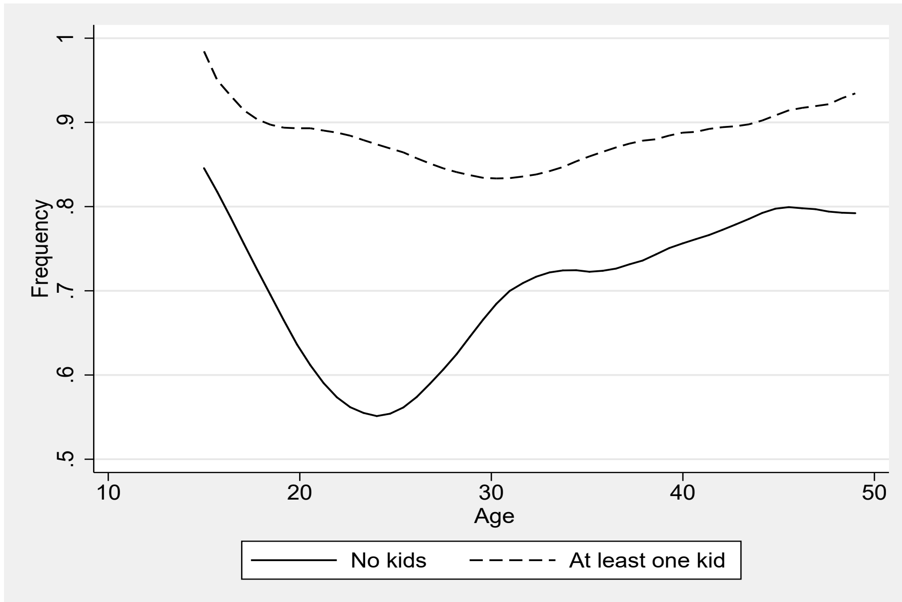


Figure 3.5: Women's self-employment, by age and level of education

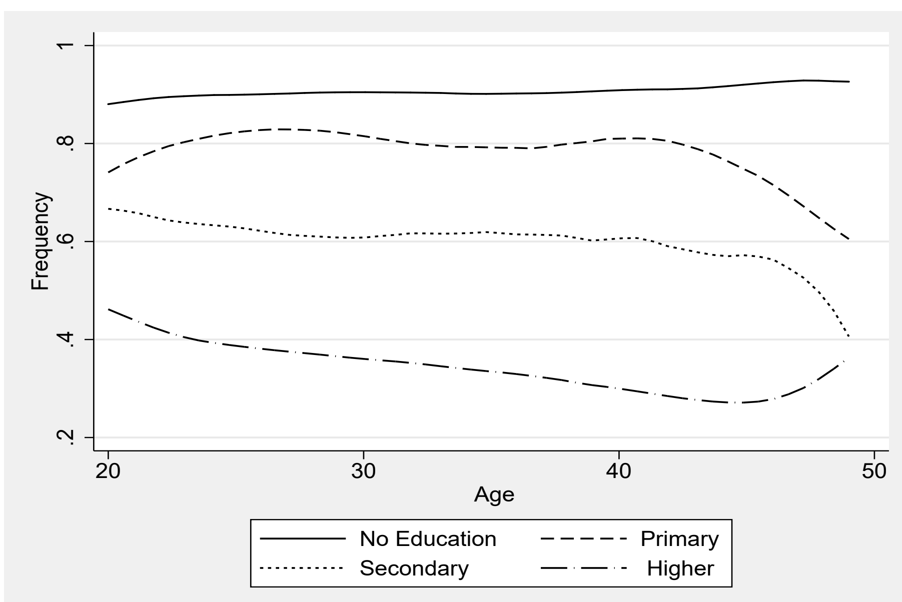
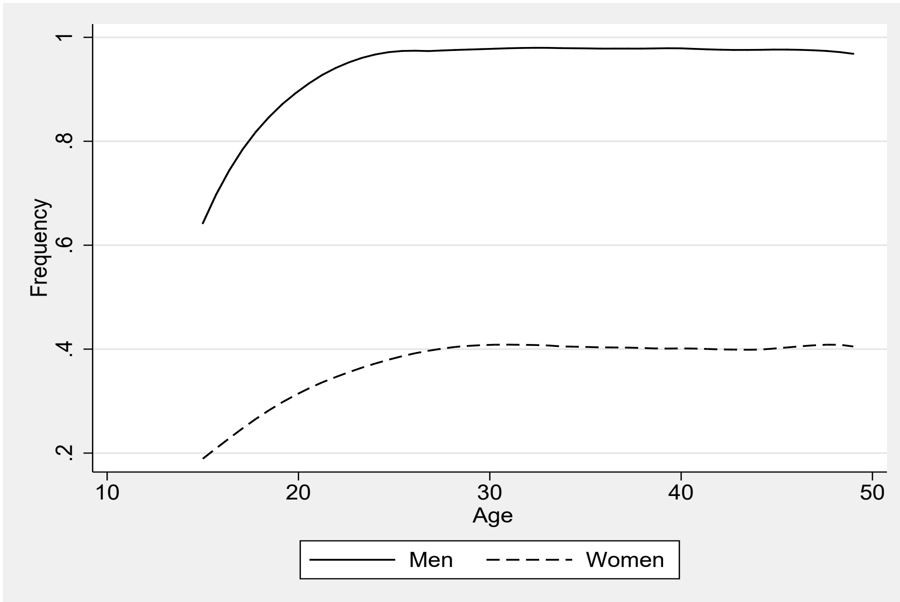


Figure 3.6: Labor force participation rates, by age and gender



[11]A woman conterparts are other women of the same age group and living in the region

Table 3.1: Summary Statistics

VARIABLES	Self-employed	Wage employed	Unemployed
Socio-demographics characteristics			
Number of preschool children	0.558	0.190	0.731
Self-reported as infertile/subfecond	0.0407	0.0344	0.0290
Education in single years of schooling	3.894	8.236	3.375
Exposure to UPE	2.892	3.753	3.708
Marital status			
Single	0.464	0.733	0.400
Married	0.536	0.267	0.600
Birth cohort			
1997-2001	0.189	0.195	0.305
1987-1996	0.323	0.433	0.333
1977-1986	0.292	0.224	0.226
1967-1976	0.197	0.147	0.136
Ethnic group			
Acholi	0.186	0.391	0.181
Aliba	0.245	0.242	0.225
Alur	0.129	0.118	0.111
Baamba	0.0771	0.0215	0.145
Baganda	0.0207	0.0275	0.0191
Others	0.342	0.199	0.318
Area of residence			
Urban	0.323	0.831	0.252
Rural	0.677	0.169	0.748
Region of residence			
Central	0.161	0.116	0.208
Eastern	0.420	0.144	0.419
Northern	0.360	0.611	0.300
Western	0.0584	0.128	0.0732
Unemployment rate among conterparts [11]	61.51	53.33	67.86

Table 3.2: Differences in individual characteristics, between working, and non working women

Variable	N	Working Women Mean	SD	Not working women Mean	SD	T-test p-value	
Education	15497	5.24	4.2	3.44	5.36	0.000	***
Preschool children	15497	0.51	0.9	0.77	0.75	0.000	***
Marital status							
Single	15497	0.44	0.47	0.34	0.5	0.000	***
Married	15497	0.56	0.47	0.66	0.5	0.000	***
Age group							
15-19	15497	0.16	0.44	0.26	0.36	0.000	***
20-29	15497	0.38	0.48	0.36	0.49	0.011	**
30-39	15497	0.3	0.43	0.25	0.46	0.000	***
40-49	15497	0.17	0.34	0.13	0.37	0.000	***
Ethnic group							
Acholi	15497	0.27	0.41	0.22	0.44	0.000	***
Aliba	15497	0.24	0.42	0.22	0.43	0.007	***
Alur	15497	0.13	0.32	0.12	0.34	0.004	***
Baamba	15497	0.05	0.32	0.12	0.21	0.000	***
Baganda	15497	0.02	0.14	0.02	0.15	0.259	
Area of residence							
urban	15497	0.48	0.44	0.26	0.5	0.000	***
rural	15497	0.52	0.44	0.74	0.5	0.000	***
Region of residence							
Central	15497	0.15	0.4	0.2	0.36	0.000	***
Eastern	15497	0.34	0.49	0.42	0.47	0.000	***
Northern	15497	0.44	0.46	0.31	0.5	0.000	***
Western	15497	0.08	0.26	0.07	0.27	0.084	*

Significance threshold: *** p<0.01, ** p<0.05, * p<0.1

Table 3.3: Presence of other adults and children during the interview about sexual activity and women's self-reported infertility status

Variable	Infertile b0	SD	Fertile b1	SD	T-Test (b0=b1) p-value
Presence of husband during interview[12]	0.13	0.31	0.11	0.34	(0.405)
Presence of other male adults during interview	0.06	0.2	0.04	0.24	(0.315)
Presence of other female adults during interview	0.1	0.28	.08	0.3	(0.477)
Presence of other adults during interview	0.1	0.29	0.09	0.3	(0.725)
Presence of children under 10 during interview	0.03	0.14	0.02	.17	(0.231)

Table 3.4: Correcting for selection and endogeneity of motherhood separately

	One-Equation model	Two-Equation model		Two-Equation model	
	(1)	(2)	(3)	(4)	(5)
	y. eqn	y. eqn	d. eqn	y. eqn	K. eqn
Education	-0.0120***	-0.00946***	0.00519***	-0.00965***	
	(0.00141)	(0.00164)	(0.00108)	(0.00169)	
Preschool children	0.00844	-0.00965	-0.0434***	0.120**	
	(0.00624)	(0.0133)	(0.00613)	(0.0577)	
Married	0.0756***	0.0150	-0.117***	0.000739	
	(0.0127)	(0.0202)	(0.0106)	(0.0376)	
Infertility status					-0.826***
					(0.0813)
Birth Cohort FE	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y
Rural dummy	Y	Y	Y	Y	Y
Ethnic group	Y	Y	Y	Y	Y
Observations	5702	12371	12371	12371	5702
R-square	0.2597				
Coefficient of correlation					
corr(y eqn., d eqn.)			0.732***		
Wald test of independence (rho = 0)					
Chi2 (1) statistic			48.62		
Prob>chi2			0.000		
F test of excluded instruments					
F(1, 5686)					106.87
Prob > F					0.0000
Underidentification test					
SW Chi-sq(1)					107.7
P-val					0.0000
Weak identification test					
Cragg-Donald Wald F statistic					106.869

Standard errors in parentheses * p<.10, ** p<.05, *** p<.01

Table 3.5: Correcting for selection and endogeneity of motherhood jointly

	Three-equation model		
	(1)	(2)	(3)
	K.eqn	y. eqn	d. eqn
Infertility status	-0.826***		
	(0.0813)		
Education		-0.00959***	0.00504***
		(0.00164)	(0.00108)
Preschool children		-0.0409	-0.0790***
		(0.0331)	(0.0144)
Married		0.0356	-0.0928***
		(0.0269)	(0.0140)
URC			0.637***
			(0.0419)
Residual_child		0.0274	0.0326***
		(0.0239)	(0.0120)
Birth Cohort FE	Y	Y	Y
Region FE	Y	Y	Y
Rural dummy	Y	Y	Y
Ethnic group	Y	Y	Y
Observations	12371	12371	12371
F-statistic (First-stage)	103.29		
Prob>chi2	0.000		
Coefficient of correlation			
corr(y eqn., d eqn.)			0.73***
Wald test of independence (rho = 0)			
Chi2 (1) statistic			47.83
Prob>chi2			0.0000

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01

Table 3.6: Correcting for selection and endogeneity of motherhood and education jointly

	Four-equation estimation			
	(1)	(2)	(3)	(4)
	E eqn.	K eqn.	y eqn.	d eqn.
Exposure to UPE	0.147***			
	(0.0102)			
Infertility status		-0.826***		
		(0.0813)		
Education			-0.0305***	-0.000388
			(0.00483)	(0.00276)
Preschool children			-0.0484	-0.0786***
			(0.0318)	(0.0145)
Married			0.000285	-0.101***
			(0.0268)	(0.0143)
Residual_education			0.0461***	0.0115**
			(0.00965)	(0.00527)
Residual_child			0.0296	0.0321***
			(0.0232)	(0.0120)
URC				0.630***
				(0.0429)
Birth Cohort FE	Y	Y	Y	Y
Region FE	Y	Y	Y	Y
Rural dummy	Y	Y	Y	Y
Ethnic group	Y	Y	Y	Y
Observations	12371	12371	12371	12371
F-statistic (First-stage)	209.22	103.29		
Prob>chi2	0.000	0.000		
Coefficient of correlation				
corr(y eqn., d eqn.)				0.767***
Wald test of independence (rho = 0)				
Chi2 (1) statistic				48.78
Prob>chi2				0.0000

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01

Table 3.7: Robustness check - Alternative definition of motherhood

	Four-equation estimation			
	(1)	(2)	(3)	(4)
	E eqn.	K eqn.	y eqn.	d eqn.
Exposure to UPE	0.147***			
	(0.0102)			
Infertility status		-0.454***		
		(0.0572)		
Education			-0.0308***	-0.00114
			(0.00482)	(0.00277)
Having preschool children			-0.0971	-0.234***
			(0.0755)	(0.0367)
Married			0.0105	-0.0561***
			(0.0327)	(0.0186)
Residual_education			0.0467***	0.0125**
			(0.00969)	(0.00527)
Residual_child			0.0509	0.103***
			(0.0408)	(0.0211)
URC				0.633***
				(0.0427)
Birth Cohort FE	Y	Y	Y	Y
Region FE	Y	Y	Y	Y
Rural dummy	Y	Y	Y	Y
Ethnic group	Y	Y	Y	Y
Observations	12371	12371	12371	12371
F-statistic (First-stage)	209.22	62.92		
Prob>chi2	0.000	0.000		
Coefficient of correlation				
corr(y eqn., d eqn.)				0.758***
				(0.0609)
Wald test of independence (rho = 0)				
Chi2 (1) statistic				47.95
Prob>chi2				0.0000

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01

Table 3.8: Robustness check - Living in Rural versus Urban area

	Four-equation estimation							
	Rural				*Urban*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	E eqn.	K eqn.	y eqn.	d eqn.	E eqn.	K eqn.	y eqn.	d eqn.
UPE	0.193***				0.106***			
	(0.00999)				(0.0236)			
Infertility		-1.061***				-0.311***		
		(0.111)				(0.0945)		
Education			-0.00640**	0.00163			-0.0163**	0.0187***
			(0.00261)	(0.00373)			(0.00674)	(0.00607)
Children			0.00773	-0.0856***			-0.0368	-0.0415
			(0.0107)	(0.0168)			(0.0403)	(0.0369)
Married			0.0255**	-0.0787***			-0.0182	-0.121***
			(0.0112)	(0.0177)			(0.0266)	(0.0259)
res			0.00393	0.0116*			0.0205	-0.0411***
			(0.00332)	(0.00617)			(0.0155)	(0.0141)
res1			0.00256	0.0449***			-0.00784	-0.00191
			(0.00846)	(0.0145)			(0.0267)	(0.0257)
URC				0.813***				0.408***
				(0.0630)				(0.0705)
Age FE	Y	Y	Y	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y	Y	Y	Y
Rural dummy	Y	Y	Y	Y	Y	Y	Y	Y
Ethnic group	Y	Y	Y	Y	Y	Y	Y	Y
Observations	8359	8359	8359	8359	4012	4012	4012	4012
F-statistic	371.29	90.56			20.10	10.81		
Prob>chi2	0.000	0.000			0.000	0.001		
corr(y , d)				-0.5104 *				0.897***
				(0.2189)				(0.0469)
Wald test								
Test Chi2(1)				3.62				36.71
Prob>chi2				0.0571				0.000

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01

Conclusion

This thesis shed light on the effects of workers' observable and unobservable characteristics such as their risk preferences, their ability, their perception bias, their education and their motherhood status on their decisions in the labor market. We contribute to the literature by highlighting the ability of experimental risk preferences measures (Chapter 1) and risk perception bias (Chapter 2) to predict contract choices and by clarifying the impact of motherhood and education on female self-employment in developing countries (Chapter 3).

The first chapter investigated the ability of risk preference-revealing-experiments to predict workers contract choices between a risky contract (piece rate) and a risk-free contract (fixed wage). Using field experiments, we estimated workers preferences using Holt and Laury's approach with low stake payoffs and high stakes payoff in order to capture the payoff scale effect on workers behaviour. We found that the aggregate distribution of risk preferences is stable across the treatments but individuals risk preferences change. Our theoretical model predicts that risk averse and more productive workers are likely to choose piece rate contracts over fixed wage contract. We tested these predictions using data from a contract choice experiment conducted with the same workers. We evaluated the ability of each measure to predict worker's contract choices between a piece rate contract and a series of fixed wage contract. We find that the risk preferences obtained using the high payoff treatment predict workers contract choices correctly while the risk preferences obtained using the low payoff treatment do not. For high-stakes lotteries experiment, more risk averse implies less likely to accept piece rate contract. Hence, increasing the stakes of the lotteries led workers to reveal their true risk preferences. In terms of implication, our results first imply that aggregate distribution of risk preference is not an effective measure of the payoff scale effect. Using the aggregated distribution of risk preferences to measure the scale effect is likely misleading.

Second, our findings imply that, when using lottery experiments to measure individuals risk preferences, the financial incentives must be high enough to incite individuals to reveal their true preferences.

The second chapter investigates the presence of risk perception bias among workers who are paid piece rates and, in case of bias, the potential of this perception bias to explain the missing risk incentive tradeoff. Using field experiments, we have measured the perception bias on the earning risk and the average earnings among workers who are paid piece rates. Then, compared the perceived distribution with the actual distribution of workers earnings. The results clearly indicate that workers underestimate the risk they face under the piece rate contract. Subsequently, we investigate the effect of this perception bias on workers choice between a piece rate contract and a fixed wage contract. The results demonstrate that perception bias is promising as an explanation for the lack of a risk-incentives tradeoff in contractual data. Underestimating risk has a significant effect on contractual choice. We find that the greater is the degree of underestimation, the more likely workers are to select piece rates. Risk averse workers will accept risky incentive contracts with a lower earnings premium than they would if they perceived earnings risk correctly. Our evidence on the interaction between risk perception and risk preferences is inconclusive and requires further study. More than 17 percent of the workers in our sample display risk loving preferences. Yet, the interaction term between risk aversion and perception bias is not significant. Future research will need more data to measure these interaction effects precisely.

The third chapter studied the role of motherhood and education to women's self-employment in a developing economy's context, using micro-level data from Uganda. Our empirical strategy conjointly address the issues of nonrandom self-selection into the paid labor force and the endogeneity of education and motherhood. We obtain consistent estimates of the effects of covariates of interests by estimating a four-equation econometric model, using a Full Information Maximum Likelihood Estimator's method combined with a control function approach for endogenous regressors. Our result contributes to the literature on female self-employment by showing that, unlike in a developed country, motherhood does not explain women over-representation in self-employment in a developing country context, where self-employment is predominantly a feature of the insecure informal economy. Instead, our results suggest that in the latter context, education reduces women's self-employment probabilities. Therefore, in a developing economy, public policy aimed at achieving gender parity in pay is likely to be more effective if it focuses on promoting female education as well as households' access to quality professional childcare services. Building women's human capital and skills through increased access to education can wean them off chronic dependence on low-paid self-employments.