Detailed Description of Methodology

First, to investigate the income differential relationship, this paper makes use of two techniques. We start with the Heckman (1979) model and estimate the conditional annual labour wage (income) amongst occupations, taking into account labour force participation. Next, we utilize quantile regressions as they compute several different regression curves corresponding to the various percentage points on the labour wage distribution, thus giving a more complete picture of the wage spectrum and enabling us to distinguish between low-value-added and high-value-added activities as implied by achieved income.

Our motivation to apply the Heckman selection model is to note that people who work are selected non-randomly from the population; we have access to annual income observations only for those who work and estimating the determinants of wages from the subpopulation who work may thus introduce a selection bias. The Heckman correction solves this potential selection bias and estimation takes place in two stages. In the first stage, we formulate a model, based on economic theory, for the probability of working as opposed to belonging to the "non- active" labour force. The canonical specification for this relationship is the initial probit equation:

$$\Pr(Employed = 1 \mid W) = \phi\,(W'\beta)$$

$$P(Employed = 1 \mid W) = P(x'_i E + x'_i X + e_i > 0 \mid W) \qquad [1.1]$$

Where *Employed* indicates employment (*Employed* = 1 if the respondent is employed, i.e. the individual has either an employment status of employer, self-employed own account individual, or paid worker), and *Employed* = 0 non-employed (the individual has the employment status of either unemployed or not in labour force), thus the variable W implies selection into occupational

categories but not income. As before, *E* is the vector that includes the different levels of education, *X* is a vector of control variables, and *e* is the error term. β is a vector of control parameters, and ϕ is the cumulative distribution function of the standard normal distribution. The estimates from this model yield results that can be used to predict the employment probability for each individual, given each individual's observable characteristics.

The Heckman procedure next corrects for self-selection by incorporating a transformation of these predicted individual probabilities as an additional explanatory variable in the second stage wage equation, which is now specified as:

$$Y_i = x_i\beta + \delta Se_i + \eta Em_i + \mu_i \qquad [1.2]$$

Where Y_i denotes logged annual income of individuals. The dummies Se_i and Em_i take the value of 1 if person *i* is a self-employed own account individual or employer respectively. The estimated coefficients $\hat{\delta}$ and $\hat{\eta}$ are interpreted as a measure of the Heckman corrected conditional earnings premium/penalty experienced by the self-employed and employers respectively, compared to paid workers. We can denote Y_i^* as the underlying wage offer, which is not observed if the respondent does not work. The conditional expectation of income given the person works is then:

$$\mathrm{E}[Y_i|x_i, W = 1] = x_i\beta + \delta Se_i + \eta Em_i + \mathrm{E}[\mu_i|x_i, W = 1]$$

$$\mathrm{E}[\mathrm{Y_i}|\mathrm{x_i}, \mathrm{X} = 1] = \mathrm{x_i}\,\beta + \delta\,\mathrm{Se_i} + \eta\,\mathrm{Em_i} + \rho\ \sigma_{\mu}\ \lambda\,(\mathrm{X}'\,\beta) \qquad [1.3]$$

Where ρ is the correlation between unobserved determinants of propensity to work E and unobserved determinants of wage offers μ_i, the variable *W* implies selection into employment, σ_u is the standard deviation of μ_i, and λ is the inverse Mills ratio evaluated at $W'\beta$. This

equation thus solves the sample selection issue that can be viewed as a form of omitted-variables bias, as conditional on x_i, δ, η and on λ as if the sample is randomly selected.

In all our estimations, Y_i denotes logged annual income, explanatory variables x_i comprise standard human capital formation literature variables (age, age squared, sector, gender, marital status, educational attainment, ability to speak a Nigerian language, region and credit constraints). The dummies Se_i and Em_i take the value of 1 if person *i* is a self-employed own account worker or employer respectively. The estimated coefficients $\hat{\delta}$ and $\hat{\eta}$ are thus interpreted as a measure of the conditional earnings premium/penalty experienced by the self-employed and employers respectively compared to paid wage earners.

This paper also makes use of quantile regressions in addition to Heckman estimates to calculate self-employment own account and employer premiums/penalties conditional on observable individual characteristics. We do this because while the Heckman selection based regression coefficients give a grand summary of averages in the distribution, quantile regressions offer additional insights since we can compute several different sets of coefficients corresponding to various percentage points specified, thus giving a more complete picture of the wage distribution (Koenker and Hallock, 2001). Thus, our quantile regression sorts employers, paid workers and self employed own account individuals into percentiles at the .05, .10, .30, .50, .70, .90, .95 and .99 positions on the income distribution and compares them to each other.

The quantile regression allows us to determine if individuals in any of the employment states are bound to experience a labour wage/income premium or penalty relative to another employment option, as evaluated at different points on the labour income distribution. Thus the expression for

any worker i at the τ^{th} quantile of the Y distribution conditional on observable characteristics can be thus expressed as:

$$F_{yi}^{-1}(\tau|x_i) = x_i\beta(\tau) + \delta(\tau)Se_i + \eta(\tau)Em_i + \mu_i, \quad \forall\tau \in [0,1] \qquad [1.4]$$

For all estimations in this paper, the base employment category was the paid worker category; the category of educational attainment that was left out of the estimations was the 'no education' category, for marriage it was 'unmarried', for region it was the 'middle-belt' of the country, for language it was those who do not speak a Nigerian language and for the location it was the rural location.

Table 1 reports the variables used in our empirical analyses and Table 2 reports the descriptive statistics of the variables. Table 3 shows how annual labour wages are distributed in the survey. Table 4 presents both the Heckman-corrected conditional labour wage premium/penalty estimates for the whole sample and quantile regression estimates based on models [1.3] and [1.4]. Figure 1 is a graph of the labour wage gap conditional on observable characteristics from estimations [1.3] and [1.4] for the whole employed sample. The horizontal y-axis represents the conditional labour wage of paid workers, while the dotted and thick black lines show the conditional income premiums/penalties of employers and self-employed own account individuals respectively. The labour wages captured on the graph represents the responses from the wage-earning sample in the survey i.e. employers, paid workers and self-employed own account individuals. In examining the impact of occupational choice on wages for men and women, we estimate equations [1.3] and [1.4] for men and women separately. These results are presented in Table 5 and Figure 2 for men and Table 6 and Figure 3 for women.

Detailed Description of Results

Table 4 presents both the Heckman-corrected conditional labour wage premium/penalty estimates for the whole sample and quantile regression estimates based on models [1.3] and [1.4]. The results indicate that self-employed own account individuals earn about 16 percent less than wage workers, while employers earn about 12.6 percent more than wage workers. The quantile regressions suggest a distinct pattern with self-employed own account individuals consistently earning less than paid employees until Q(.80). Crucially, the labour wage differential between paid workers and self-employed own account workers steadily increases at the lower end of the income spectrum until Q(.20), where the minimum wage of ₦18,000 is. After this point, the labour wage differential starts to reduce steadily with the differences in earnings disappearing at the top most quantiles. Throughout the earnings distribution, employers have the highest earnings.

Figure 1 is a graph of the labour wage gap conditional on observable characteristics from estimations [1.3] and [1.4] for the whole employed sample. The horizontal y-axis represents the conditional labour wage of paid workers, while the dotted and thick black lines show the conditional income premiums/penalties of employers and self-employed own account individuals respectively. The labour wages captured on the graph represents the responses from the wage-earning sample in the survey i.e. employers, paid workers and self-employed own account individuals and ranges from ₦ 3,000 to ₦ 180,000 monthly. The results in Figure 1 and Table 4 suggest that self-employed individuals are a heterogeneous group, with employers having significantly more economic returns than self-employed own account individuals.

Employers have significant earnings premiums and do better than paid workers and self-employed own account individuals throughout the income distribution, and hence we can conjecture they are mostly engaged in pulled (opportunity-driven) self-employment if only monetary indicators are considered. Self-employed own account individuals are worse off in terms of labour wage until about the 80th quantile, however from that point onwards they begin to have earnings comparable with employers and paid workers.

In examining the impact of occupational choice on wages for men and women, we estimate equations [1.3] and [1.4] for men and women separately. The results are presented in Table 5 and Figure 2 for men and Table 6 and Figure 3 for women. The results suggest that both male and female employers experience income premiums and do better than paid workers and self-employed own account individuals for most part of the distribution until the upper quantiles. The labour wage penalty for self-employed "own-account" workers occurs up to a certain point on the labour wage distribution and then starts to improve. This cut-off point is the 10th percentile for men, and interestingly a higher threshold of the 20th percentile for women, consistent with what we already discussed regarding minimum wage; as self-employed individuals are not guaranteed minimum wages, they appear to be worse off towards the lower end of the distribution.

Our results in sum show that the labour wage penalty for self-employed own account individuals is found towards the lower end of the distribution where paid workers enjoy some minimum wage guarantee; this amount has been reviewed several times but is currently ₦18,000 monthly. The trend is for the labour wage penalty to progressively increase up to the minimum wage level point between the 10th and 20th percentiles and then improve later as this effect wears off. This

finding is significant and robust, and in the absence of other economic reasoning, this may indicate two things: (1) that potential workers who cannot receive employment at minimum wage are displaced towards low income self-employment as they cannot legally be hired below minimum wage, (2) that these individuals are special in the sense that they cannot even attain a job that provides the minimum wage, which is subsequently reflected in the fact that they are penalized.

As the World Bank, Nigerian Bureau of Statistics and other data providing bodies have highlighted significant regional differences in Nigeria, we examine the robustness of the results for different regions (also included in our regressions). Thus, we recreate the estimations [1.3] and [1.4] for each of the four regions in Nigeria (the North, Mid-belt, South-East and South-West). The results are robust, and consistently show that self-employed own account individuals are worse off in terms of labour wages in all the four regions and that due to the fact that self-employed individuals are not guaranteed minimum wages in the manner already described.

Results from our classification tests in Table 7 imply that women need higher educational attainments than men to: (1) become employers and (2) be in paid employment. Our results from the Heckman estimation in Table 4 indicate that men earn on average 25 percent more than women. From our Quantile regressions, we also find that these gender differences can be found throughout the income distribution but gets progressively worse towards the higher end of the distribution; men earn on average 13 percent more than women at the 5^{th} quantile to 40 percent at the 99^{th} quantile. These results are all significant at the 1 percent level.

Tables and Figures

Table 1. Variables Used in Empirical Estimations – Descriptive Statistics are presented in Table 2

Variable	What it Measures	Methodology
Employment Status	Employment Status	Employer = 1 Self-Employed (Own Account) =2 Paid Worker = 3 Unemployed = 4 Not in Labour Force = 5
Employed	Being in Employment	Dummy (1/0) [Employed = 1 if Employer, Self-Employed (Own Account) or Paid Worker] [Non-Worker = 0 if Unemployed or Not in Labour Force]
Sex	Male or Female	Dummy (1/0) [Male = 1] [Female = 0]
Age in Years	Age in years	Age in Years
Age Square	Age Squared	Age Squared
Sector	Urban or Rural Residence	Dummy (1/0) [Urban = 1] [Rural = 0]
Marital Status	Marital Status	Dummy (1/0) [Married] [Not Married = 0]
Religion	Religion	Dummy (1/0) for 4 religions: [Christian, Muslim & Others]
Educational Attainment	Educational attainment	Dummy (1/0) for 5 categories: [No Ed = No Education] [Lo Ed = Low Education (primary to junior secondary)] [Mid Ed = Medium Education (senior secondary to 'O' level)] [High Ed = High Education (BSc and equivalent)] Very High Ed = Very High Education (Masters to Doctorate)]
Region	Region of the country	Dummy (1/0) for 4 regions: [South-East = South East] [Mid-Belt = Middle Belt] [South-West = South West] [North = North]
House or Land	A proxy for source of collateral for bank loan	Dummy (1/0) measuring if the Individual owns a Plot of Land or House: [Owns = 1] [Does not own = 0]
Local Language	Ability to speak or/and write a Nigerian Language	Dummy (1/0) measuring if the Individual can speak/write a Nigerian Language: [Can Speak/Write = 1] [Cannot Speak/Write = 0]
Log (Annual Income)	Log of Annual Income	Log of Annual Income

Table 2. Descriptive Statistics

Variable	Whole Sample Mean (Std Dev)	Male Mean (Std Dev)	Female Mean (Std Dev)
Employment			
Employer	3.3%	4.4%	1.85%
Self-Employed (Own Account)	35.64%	41.12%	28.84%
Paid Worker	19.94%	25.71%	12.75%
Unemployed	29.73%	20.3%	41.55%
Non Labour Force	11.39%	8.5%	15.01
Education			
No Ed	4.4%	3.8%	5.2%
Low Ed	31.5%	33.5%	29.1%
Mid Ed	32.4%	37.5%	26.1%
High Ed	6.8%	8.6%	4.7%
Very high Ed	1.7%	2.1%	0.7%
Unspecified	23.2%	14.5%	34.2%
Demographic			
Age in years	34.32	35.14	33.3
	(13.165)	(13.114)	(13.159)
Married	55.2%	50.1%	61.8%
Christian	63.4%	65.8%	60.4%
Muslim	35.3%	32.9%	38.3%
Geographic			
Urban	42.8%	45.3%	39.7%
Rural	57.2%	54.7%	60.3%
Region			
South-East	34.2%	35%	33.2%
South-West	24.8%	25.9%	23.4%
Mid-Belt	18.5%	18.6%	18.4%
North	22.4%	20.5%	25%
Further Controls			
House or Land	10.1%	08.9%	11.8%
Local Language	67.5%	74.5%	59.0%
Log of Annual Income	6.973	8.511	5.057
	(5.8607)	(5.4462)	(5.7944)
N	18,397	10,206	8,191

Table 3. Mean Annual Labour Income of Employed Sample by Employment Type, Gender and Region

		Male		Female	
		Mean Wage ₦[1] (Std Err)	N	Mean Wage ₦ (Std Err)	N
North					
	Self-Employed (O.A)[2]	185,965.8 (9,644.5)	751	*219,091 ** (24,881.5)	140
	Paid Worker	281,630.5 (13,692.16)	621	191,768.2 (10,059.76)	137
	Employer	311,048.9 (28,595.95)	165	*412,033.3 ** (78,699.14)	24
Mid-Belt					
	Self-Employed (O.A)	206,539 (10,299.61)	703	140,146.8 (12,662.84)	313
	Paid Worker	238,019.9 (10,548.42)	507	195,603.8 (9,299.89)	153
	Employer	266,039.3 (24,226.71)	125	*293,468.8 ** (52,992.66)	32
South-East					
	Self-Employed (O.A)	208,569.5 (5,471.77)	1,650	133,462.6 (5,250.546)	1,010
	Paid Worker	260,804.9 (12,123.49)	712	222,930.6 (10,861.79)	389
	Employer	343,745.2 (29,518.05)	86	171,220 (12,743.83)	47
South-West					
	Self-Employed (O.A)	184,967.1 (6,752.93)	1093	129,438.1 (6,706.68)	898
	Paid Worker	225,743.8 (7,332.42)	784	205,809.8 (10,436.51)	366
	Employer	213,274.6 (16,228.51)	77	*358,962.9 ** (57,584.58)	49
Entire Sample					
	Self-Employed (O.A)	198,038.1 (3,700.58)	4,197	137,895.5 (4,086.32)	2,361
	Paid Worker	250,855.5 (5,515.434)	2,624	208,847.8 (5,776.33)	1,045
	Employer	288,216.8 (13,973.58)	453	*295,502.1 ** (26,074.22)	152
	Gender Total	*222,707.3 (3,065.33)*	*7,274*	*165,467.6 (3,458.49)*	*3,558*
	Survey Total	*194,087.45 (3,261.91)*			
	N	10,832	7,274		3,558

[1] Conversion rate was about 158 ₦ ≈ $1 during the writing period.
[2] "Own Account" workers represent the Self-Employed.

Table 4. Labour Income Premiums and Penalties, Estimations [1.4] and [1.5] - Whole Sample

Variables	Heckman Estimation	Quantile Estimation							
		Q(.05)	Q(.10)	Q(.30)	Q(.50)	Q(.70)	Q(.90)	Q(.95)	Q(.99)
Employer	0.126***	0.007	0.065	0.100***	0.109**	0.070**	0.286***	0.316**	0.057
	(0.033)	(0.093)	(0.082)	(0.031)	(0.055)	(0.028)	(0.071)	(0.146)	(0.148)
S.E (O.A)	-0.160***	-0.156***	-0.273***	-0.242***	-0.206***	-0.095***	0.076	0.076	0.182
	(0.017)	(0.037)	(0.029)	(0.021)	(0.023)	(0.019)	(0.051)	(0.053)	(0.108)
Age in Years	0.033***	0.035***	0.037***	0.041***	0.031***	0.033***	-0.028**	0.001	-0.038
	(0.010)	(0.006)	(0.008)	(0.004)	(0.006)	(0.007)	(0.014)	(0.018)	(0.031)
Age Square	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	0.000**	0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban	0.123***	0.113***	0.100***	0.151***	0.141***	0.109***	0.148***	0.033	0.091
	(0.017)	(0.018)	(0.027)	(0.016)	(0.025)	(0.026)	(0.040)	(0.070)	(0.112)
Unspecified	0.014	0.041	-0.002	-0.054	-0.096**	-0.004	0.284***	0.791***	0.353
	(0.041)	(0.047)	(0.063)	(0.057)	(0.045)	(0.067)	(0.100)	(0.100)	(0.332)
Low Ed	0.207***	0.157***	0.214***	0.247***	0.146***	0.194***	0.325***	0.754***	-0.041
	(0.039)	(0.046)	(0.063)	(0.048)	(0.046)	(0.060)	(0.053)	(0.101)	(0.426)
Mid Ed	0.344***	0.256***	0.324***	0.392***	0.316***	0.356***	0.453***	0.744***	0.003
	(0.040)	(0.041)	(0.063)	(0.050)	(0.051)	(0.059)	(0.072)	(0.086)	(0.432)
High Ed	0.783***	0.806***	0.945***	0.922***	0.780***	0.728***	0.793***	0.988***	-0.226
	(0.046)	(0.085)	(0.077)	(0.038)	(0.043)	(0.064)	(0.089)	(0.101)	(0.420)
Very high Ed	0.999***	0.286	1.105***	1.213***	0.980***	1.215***	1.370***	1.311***	0.178
	(0.062)	(0.233)	(0.378)	(0.067)	(0.062)	(0.078)	(0.108)	(0.084)	(0.455)
Male	0.251***	0.131***	0.150***	0.233***	0.238***	0.240***	0.272***	0.275***	0.405***
	(0.023)	(0.023)	(0.024)	(0.021)	(0.019)	(0.017)	(0.047)	(0.038)	(0.065)
Married	-0.002	0.013	-0.008	-0.020	0.013	0.016	0.053	-0.028	-0.184*
	(0.017)	(0.023)	(0.025)	(0.017)	(0.018)	(0.021)	(0.039)	(0.048)	(0.110)
Christian	0.053	0.010	-0.028	-0.019	0.049	0.031	0.297**	0.277	-0.393
	(0.069)	(0.076)	(0.090)	(0.064)	(0.108)	(0.115)	(0.126)	(0.466)	(0.496)
Muslim	0.032	0.063	-0.004	-0.018	-0.001	-0.009	0.267*	0.228	-0.515
	(0.070)	(0.074)	(0.096)	(0.068)	(0.111)	(0.106)	(0.138)	(0.480)	(0.498)
House or Land	-0.087***	-0.015	-0.059***	-0.117***	-0.069***	-0.059	-0.019	0.000	-0.000
	(0.027)	(0.020)	(0.016)	(0.031)	(0.026)	(0.041)	(0.064)	(0.069)	(0.102)
Local Language	-0.033	-0.006	-0.036	-0.098***	-0.034	0.029	-0.024	-0.045	0.087
	(0.021)	(0.033)	(0.025)	(0.023)	(0.024)	(0.028)	(0.049)	(0.041)	(0.119)
South-East	0.092***	0.139***	0.104***	0.117***	0.081***	0.054*	0.137**	0.093	0.109
	(0.023)	(0.031)	(0.035)	(0.025)	(0.018)	(0.033)	(0.062)	(0.121)	(0.128)
South-West	-0.099***	-0.058**	-0.095***	-0.047**	-0.110***	-0.130***	-0.196***	-0.210*	-0.180*
	(0.024)	(0.028)	(0.031)	(0.023)	(0.032)	(0.028)	(0.057)	(0.114)	(0.096)
North	0.039	0.009	-0.025	0.021	-0.024	-0.023	0.150	0.519***	0.156
	(0.026)	(0.025)	(0.032)	(0.025)	(0.020)	(0.031)	(0.094)	(0.094)	(0.126)
Cons	10.602***	9.696***	9.971***	10.202***	10.679***	10.833***	12.246***	11.756***	14.949***
	(0.254)	(0.145)	(0.148)	(0.121)	(0.147)	(0.126)	(0.358)	(0.564)	(0.999)
Sigma R^2	0.046	0.0796	0.1080	0.1545	0.1348	0.1071	0.0627	0.0537	0.0348

* p<0.1; ** p<0.05; *** p<0.01, Dependent Variable is Log of Annual Wage

Table 5. Labour Income Premiums and Penalties, Estimations [1.4] and [1.5] - Male Sample

Variables	Heckman Estimation	Quantile Estimation							
		Q(.05)	Q(.10)	Q(.30)	Q(.50)	Q(.70)	Q(.90)	Q(.95)	Q(.99)
Employer	0.115***	0.056	0.144*	0.099***	0.130*	0.131***	0.155**	0.148	0.199
	(0.038)	(0.051)	(0.086)	(0.027)	(0.072)	(0.036)	(0.064)	(0.147)	(0.192)
S.E (O.A)	-0.086***	-0.101**	-0.198***	-0.164***	-0.121***	0.029	0.108**	0.143	0.183
	(0.021)	(0.045)	(0.038)	(0.025)	(0.030)	(0.019)	(0.046)	(0.091)	(0.106)
Age in Years	-0.015	0.044***	0.046***	0.029***	0.012*	0.023***	-0.039***	-0.012	-0.092**
	(0.011)	(0.009)	(0.009)	(0.008)	(0.006)	(0.006)	(0.015)	(0.025)	(0.038)
Age Square	0.000	-0.000***	-0.001***	-0.000***	-0.000	-0.000**	0.000***	0.000	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban	0.199***	0.191***	0.231***	0.179***	0.189***	0.198***	0.192***	0.235***	0.153
	(0.020)	(0.035)	(0.035)	(0.030)	(0.026)	(0.031)	(0.037)	(0.067)	(0.104)
Unspecified	-0.111**	-0.105*	-0.136*	-0.254***	-0.176***	-0.222**	0.222**	0.741***	-0.029
	(0.053)	(0.056)	(0.081)	(0.066)	(0.063)	(0.113)	(0.087)	(0.192)	(0.540)
Low Ed	0.097*	0.067	0.110	0.086	0.063	0.019	0.243**	0.628***	-0.150
	(0.050)	(0.068)	(0.076)	(0.075)	(0.060)	(0.110)	(0.094)	(0.175)	(0.511)
Mid Ed	0.231***	0.137	0.222***	0.229***	0.212***	0.199*	0.338***	0.583***	-0.095
	(0.050)	(0.087)	(0.072)	(0.071)	(0.058)	(0.104)	(0.056)	(0.166)	(0.473)
High Ed	0.717***	0.934***	0.859***	0.740***	0.659***	0.577***	0.698***	0.816***	-0.398
	(0.057)	(0.120)	(0.066)	(0.084)	(0.059)	(0.119)	(0.077)	(0.151)	(0.520)
Very High Ed	0.850***	0.151	0.074	0.903***	0.911***	0.989***	1.412***	1.438***	0.090
	(0.075)	(0.183)	(0.525)	(0.095)	(0.085)	(0.137)	(0.153)	(0.181)	(0.479)
Married	-0.005	-0.004	0.007	0.028	0.039	0.045	0.101*	-0.014	0.087
	(0.025)	(0.045)	(0.026)	(0.032)	(0.024)	(0.033)	(0.057)	(0.071)	(0.161)
Christian	0.014	-0.161***	-0.050	0.045	0.060	0.047	0.314	-0.040	-0.369
	(0.088)	(0.052)	(0.086)	(0.119)	(0.104)	(0.158)	(0.282)	(0.576)	(0.272)
Muslim	0.014	-0.093	0.008	0.085	0.057	-0.005	0.256	-0.182	-0.623**
	(0.090)	(0.062)	(0.117)	(0.128)	(0.106)	(0.162)	(0.255)	(0.560)	(0.245)
House or Land	-0.050	-0.033	-0.036	-0.102***	-0.051*	-0.030	-0.000	0.069	0.025
	(0.033)	(0.028)	(0.030)	(0.039)	(0.030)	(0.029)	(0.080)	(0.113)	(0.136)
Local Language	0.002	-0.038	-0.030	-0.068***	0.048*	-0.007	0.017	-0.029	-0.029
	(0.026)	(0.058)	(0.031)	(0.025)	(0.025)	(0.034)	(0.059)	(0.061)	(0.227)
South-East	0.139***	0.156***	0.157***	0.185***	0.179***	0.063**	0.136**	-0.024	0.116
	(0.028)	(0.036)	(0.036)	(0.041)	(0.031)	(0.027)	(0.064)	(0.105)	(0.127)
South-West	-0.137***	-0.067	-0.095**	-0.041	-0.114***	-0.109***	-0.197***	-0.466***	-0.087
	(0.028)	(0.057)	(0.042)	(0.038)	(0.039)	(0.029)	(0.048)	(0.134)	(0.099)
North	0.001	-0.004	-0.052	-0.001	-0.037	-0.025	0.136*	0.311***	0.389**
	(0.030)	(0.042)	(0.047)	(0.041)	(0.034)	(0.041)	(0.071)	(0.089)	(0.191)
Cons	12.001***	9.835***	9.948***	10.642***	11.233***	11.349***	12.796***	12.862***	16.641***
	(0.277)	(0.164)	(0.264)	(0.224)	(0.150)	(0.261)	(0.369)	(0.857)	(0.859)
Sigma R^2	0.7515	0.0936	0.1173	0.1295	0.1119	0.0931	0.0574	0.0447	0.0397

* p<0.1; ** p<0.05; *** p<0.01, Dependent Variable is Log of Annual Wage

Table 6. Labour Income Premiums and Penalties, Estimations [1.4] and [1.5] – Female Sample

Variables	Heckman Estimation	Quantile Estimation							
		Q(.05)	Q(.10)	Q(.30)	Q(.50)	Q(.70)	Q(.90)	Q(.95)	Q(.99)
Employer	0.172***	0.000	0.041	-0.001	0.044	0.056	0.998***	0.636***	0.026
	(0.062)	(0.144)	(0.103)	(0.057)	(0.106)	(0.084)	(0.215)	(0.185)	(0.229)
S.E (O.A)	-0.326***	-0.227***	-0.326***	-0.452***	-0.382***	-0.303***	-0.180**	-0.079	0.016
	(0.031)	(0.049)	(0.031)	(0.029)	(0.035)	(0.060)	(0.078)	(0.084)	(0.175)
Age in Years	0.079***	0.013	0.024***	0.041***	0.038***	0.033***	0.001	-0.028	0.032
	(0.020)	(0.008)	(0.008)	(0.009)	(0.011)	(0.012)	(0.025)	(0.030)	(0.067)
Age Square	-0.001***	-0.000	-0.000**	-0.000***	-0.000**	-0.000**	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Urban	-0.016	-0.007	-0.024	0.021	0.009	-0.022	-0.135	-0.192**	-0.147
	(0.032)	(0.031)	(0.027)	(0.025)	(0.033)	(0.054)	(0.086)	(0.075)	(0.229)
Unspecified	0.143**	0.067**	0.090	0.058	0.020	0.208**	0.515***	0.765***	1.098
	(0.066)	(0.029)	(0.059)	(0.074)	(0.090)	(0.096)	(0.120)	(0.222)	(0.713)
Low Ed	0.369***	0.218***	0.303***	0.309***	0.238***	0.309***	0.452***	0.790***	0.827
	(0.063)	(0.057)	(0.074)	(0.077)	(0.084)	(0.082)	(0.119)	(0.157)	(0.660)
Mid Ed	0.534***	0.289***	0.433***	0.470***	0.500***	0.482***	0.562***	0.750***	0.873
	(0.069)	(0.058)	(0.056)	(0.081)	(0.087)	(0.083)	(0.143)	(0.158)	(0.676)
High Ed	0.962***	0.707***	0.723***	1.063***	0.963***	0.884***	1.064***	1.227***	1.127
	(0.081)	(0.067)	(0.110)	(0.090)	(0.089)	(0.072)	(0.151)	(0.143)	(0.730)
Very High Ed	1.393***	0.216	1.690***	1.382***	1.544***	1.440***	1.238***	1.310***	0.882
	(0.122)	(0.735)	(0.224)	(0.129)	(0.247)	(0.107)	(0.159)	(0.196)	(0.737)
Married	0.022	0.021	0.033	-0.016	-0.013	0.003	0.159***	0.122	-0.055
	(0.028)	(0.036)	(0.038)	(0.034)	(0.039)	(0.033)	(0.050)	(0.095)	(0.220)
Christian	0.083	0.057	0.205**	-0.069	-0.061	-0.059	0.331	-0.264	0.784**
	(0.108)	(0.055)	(0.101)	(0.198)	(0.178)	(0.114)	(0.434)	(0.480)	(0.330)
Muslim	-0.028	0.077	0.158*	-0.129	-0.170	-0.145	0.197	-0.109	0.828**
	(0.113)	(0.055)	(0.091)	(0.210)	(0.177)	(0.107)	(0.454)	(0.494)	(0.324)
House or Land	-0.135***	-0.007	-0.038	-0.077*	-0.082*	-0.099*	-0.179*	-0.129	-0.410
	(0.048)	(0.048)	(0.042)	(0.045)	(0.049)	(0.059)	(0.099)	(0.178)	(0.253)
Local Language	-0.092***	-0.064	-0.082**	-0.167***	-0.151***	-0.056	0.026	-0.027	-0.136
	(0.034)	(0.039)	(0.041)	(0.036)	(0.050)	(0.043)	(0.088)	(0.100)	(0.160)
South-East	0.081*	0.091**	0.004	-0.003	-0.092**	-0.073*	0.259***	0.319***	-0.141
	(0.048)	(0.036)	(0.047)	(0.036)	(0.039)	(0.042)	(0.087)	(0.117)	(0.283)
South-West	0.068	-0.027	-0.081**	-0.041	-0.080	-0.087*	0.044	0.221	-0.135
	(0.055)	(0.043)	(0.040)	(0.046)	(0.051)	(0.046)	(0.091)	(0.175)	(0.286)
North	0.082	0.132***	0.118**	0.167***	0.100**	0.073	0.397*	0.944***	0.119
	(0.068)	(0.046)	(0.055)	(0.063)	(0.043)	(0.050)	(0.241)	(0.288)	(0.390)
Cons	9.381***	10.230***	10.034***	10.456***	10.787***	11.065***	11.430***	12.571***	11.455***
	(0.530)	(0.170)	(0.199)	(0.315)	(0.291)	(0.295)	(0.568)	(0.465)	(1.772)
Sigma R^2	0.7369	0.0705	0.1013	0.1578	0.1634	0.1331	0.0929	0.0885	0.0644

* p<0.1; ** p<0.05; *** p<0.01, Dependent Variable is Log of Annual Wage

Table 7. Results of Multinomial Probit Selection Estimation [Marginal Effects]

	Male				Female			
Independent Variables:	Employer	Self-Employed (O.A)	Paid Work	Non-Employed[3]	Employer	Self-Employed (O.A)	Paid Work	Non-Employed
Age in Years	0.0139***	0.0719***	0.0712***	-0.157***	0.00453***	0.0741***	0.0415***	-0.120***
	(0.00214)	(0.00722)	(0.00718)	(0.00802)	(0.000713)	(0.00676)	(0.00313)	(0.00770)
Age (Squared)	-0.00015***	-0.000818***	-0.00076***	0.00173***	-0.0001***	-0.00076***	-0.0005***	0.00127***
	(-0.00002)	(-0.00008)	(-0.00008)	(0.000102)	(-0.00009)	(-0.00008)	(-0.00004)	(0.000101)
Urban	-0.0129	0.00232	0.0351	-0.0246	-0.00006	0.0486*	-0.000887	-0.0477
	(0.00855)	(0.0262)	(0.0296)	(0.0383)	(0.00281)	(0.0261)	(0.0120)	(0.0291)
Unspecified Ed	-0.0224***	-0.00167	-0.0720	0.0961	0.00524	-0.0994**	-0.0177	0.112*
	(0.00807)	(0.0767)	(0.0642)	(0.0908)	(0.0113)	(0.0489)	(0.0288)	(0.0600)
Low Ed	-0.00761	-0.148**	0.158**	-0.00260	0.0384	-0.0503	0.0598	-0.0479
	(0.0143)	(0.0597)	(0.0726)	(0.0639)	(0.0262)	(0.0475)	(0.0436)	(0.0677)
Mid Ed	-0.00122	-0.306***	0.230***	0.0776	0.0455**	-0.135***	0.225***	-0.135*
	(0.0160)	(0.0604)	(0.0673)	(0.0634)	(0.0680)	(0.0415)	(0.0654)	(0.0745)
High Ed	0.000642	-0.449***	0.319***	0.129	0.0735*	-0.222***	0.456***	-0.308***
	(0.0176)	(0.0276)	(0.0822)	(0.0863)	(0.0549)	(0.0217)	(0.107)	(0.107)
Very High Ed	0.0282	-0.409***	0.443***	-0.0621	0.123	-0.245***	0.634***	-0.512***
	(0.0370)	(0.0233)	(0.0867)	(0.0919)	(0.111)	(0.0166)	(0.141)	(0.121)
Married	0.0183**	0.244***	0.0992***	-0.361***	-0.00305	0.114***	0.00823	-0.119***
	(0.00715)	(0.0345)	(0.0329)	(0.0304)	(0.00395)	(0.0238)	(0.0139)	(0.0300)
House or Land	0.0104	0.00257	-0.0701***	0.0572	-0.00605**	-0.0156	-0.00127	0.0229
	(0.0103)	(0.0315)	(0.0263)	(0.0393)	(0.00257)	(0.0254)	(0.0150)	(0.0291)
Local Language	0.00279	0.0158	-0.0230	0.00439	0.00615	0.0352	0.0208	-0.0622
	(0.00709)	(0.0376)	(0.0337)	(0.0359)	(0.00424)	(0.0324)	(0.0166)	(0.0381)
South-East	-0.0255***	0.0729*	-0.0865***	0.0391	-0.00755**	0.198***	0.0385*	-0.229***
	(0.00776)	(0.0381)	(0.0327)	(0.0429)	(0.00323)	(0.0419)	(0.0218)	(0.0392)
South-West	-0.0147	0.0817**	-0.0720***	0.00503	-0.00390	0.232***	0.0400*	-0.269***
	(0.00898)	(0.0348)	(0.0275)	(0.0377)	(0.00464)	(0.0342)	(0.0232)	(0.0329)
North	0.0218	-0.0504	0.0167	0.0118	-0.00519	-0.246***	-0.0228	0.274***
	(0.0158)	(0.0361)	(0.0306)	(0.0364)	(0.00416)	(0.0228)	(0.0193)	(0.0280)
Log-pseudo likelihood	-54081132	-54081612	-54282972	-56425771	-43623821	-43624372	-43684483	-45092930
Frequency	453	4,197	2,624	2,932	152	2,361	1,045	4,633
Wald (Prob > chi2)	7281.91***	7281.91***	7281.91***	7281.91***	3108.79***	3108.79***	3108.79***	3108.79***

[3] The results suggest that there are no significant differences in educational attainments between unemployed individuals and individuals not in the labour force and these two groups are combined into the "non-employed" sample in this estimation.

Figure 1. Heckman and Quantile Labour Income Premiums and Penalties – Total Sample

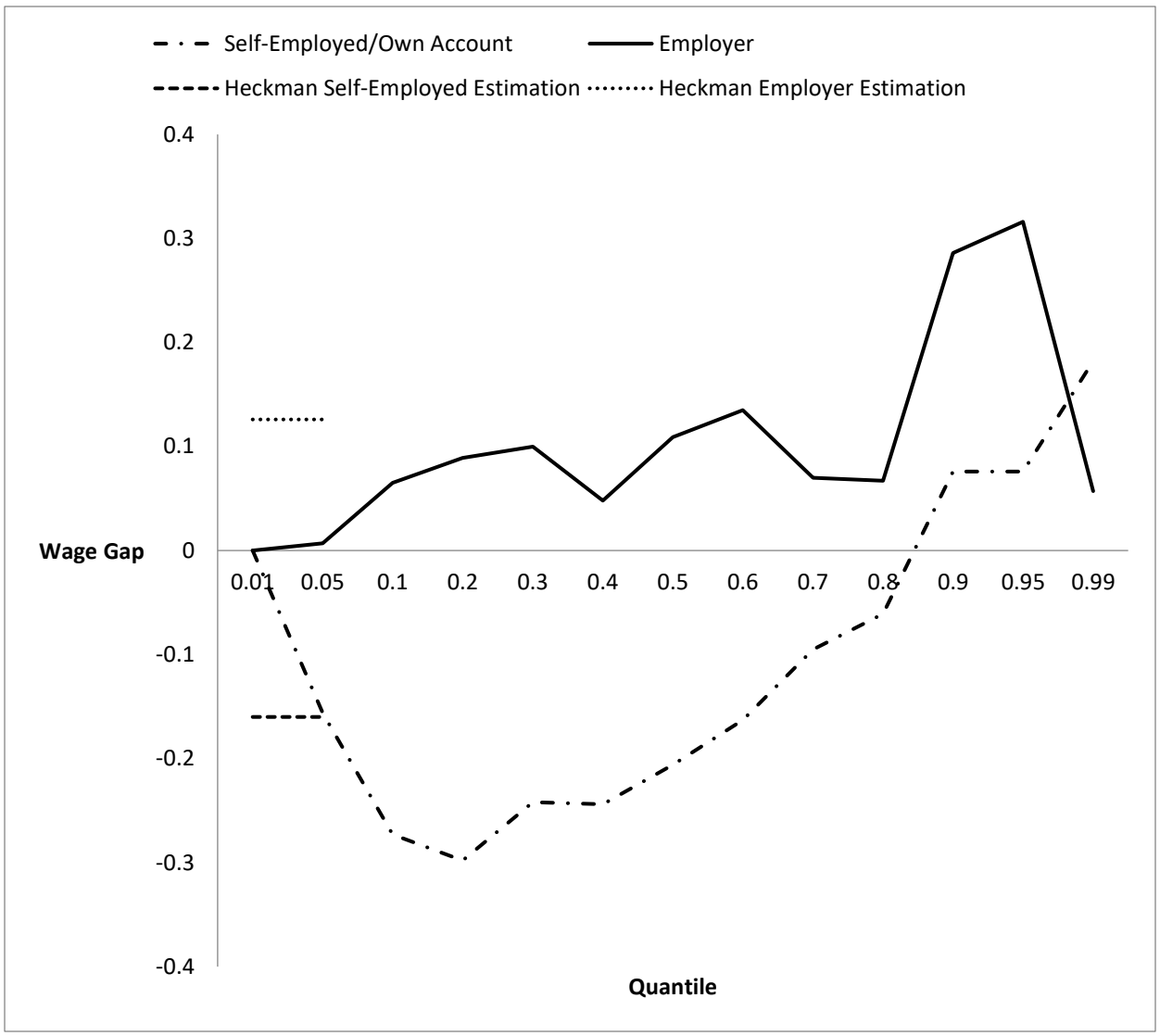


Figure 2. Heckman and Quantile Labour Income Premiums and Penalties – Male Sample

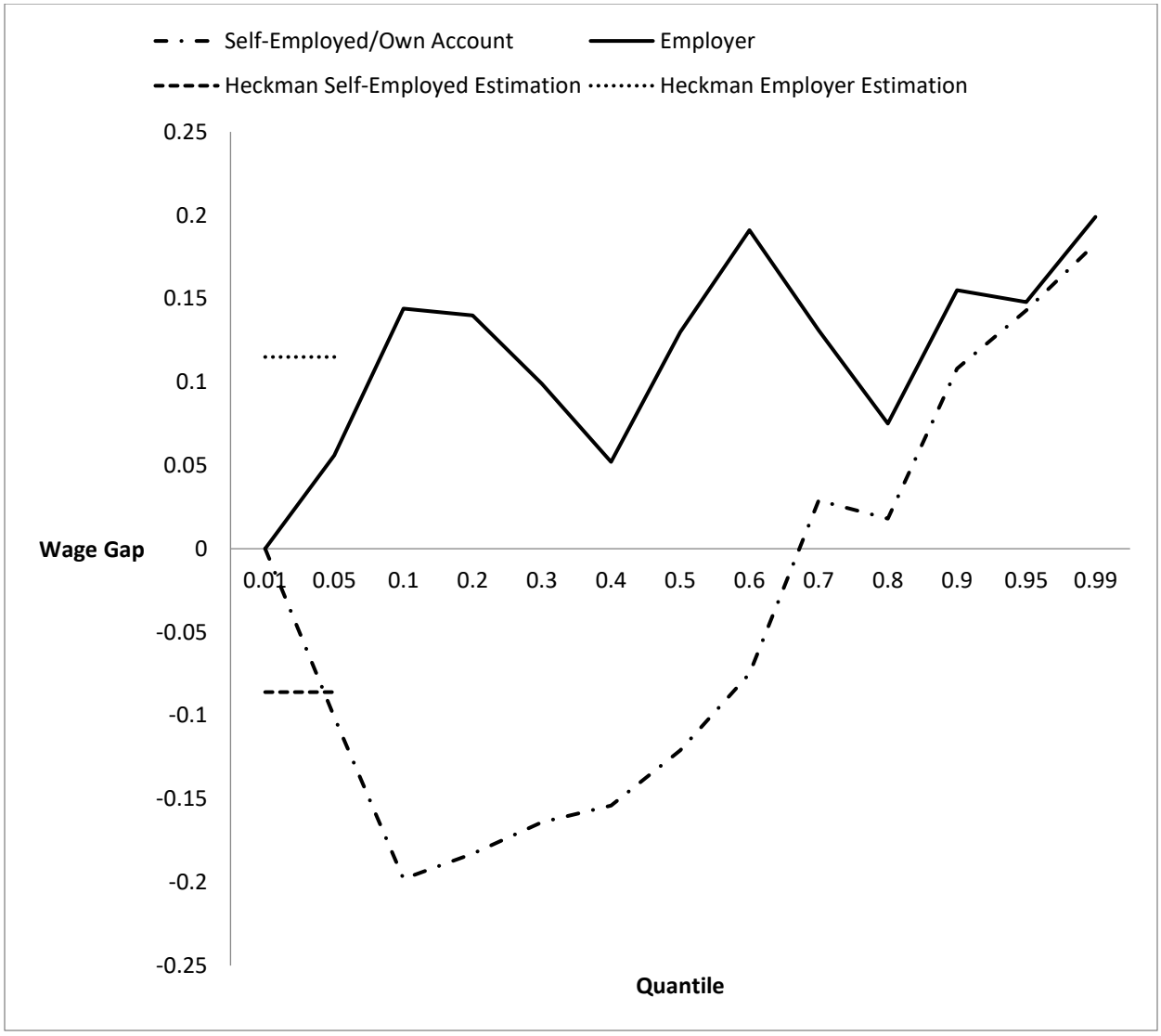
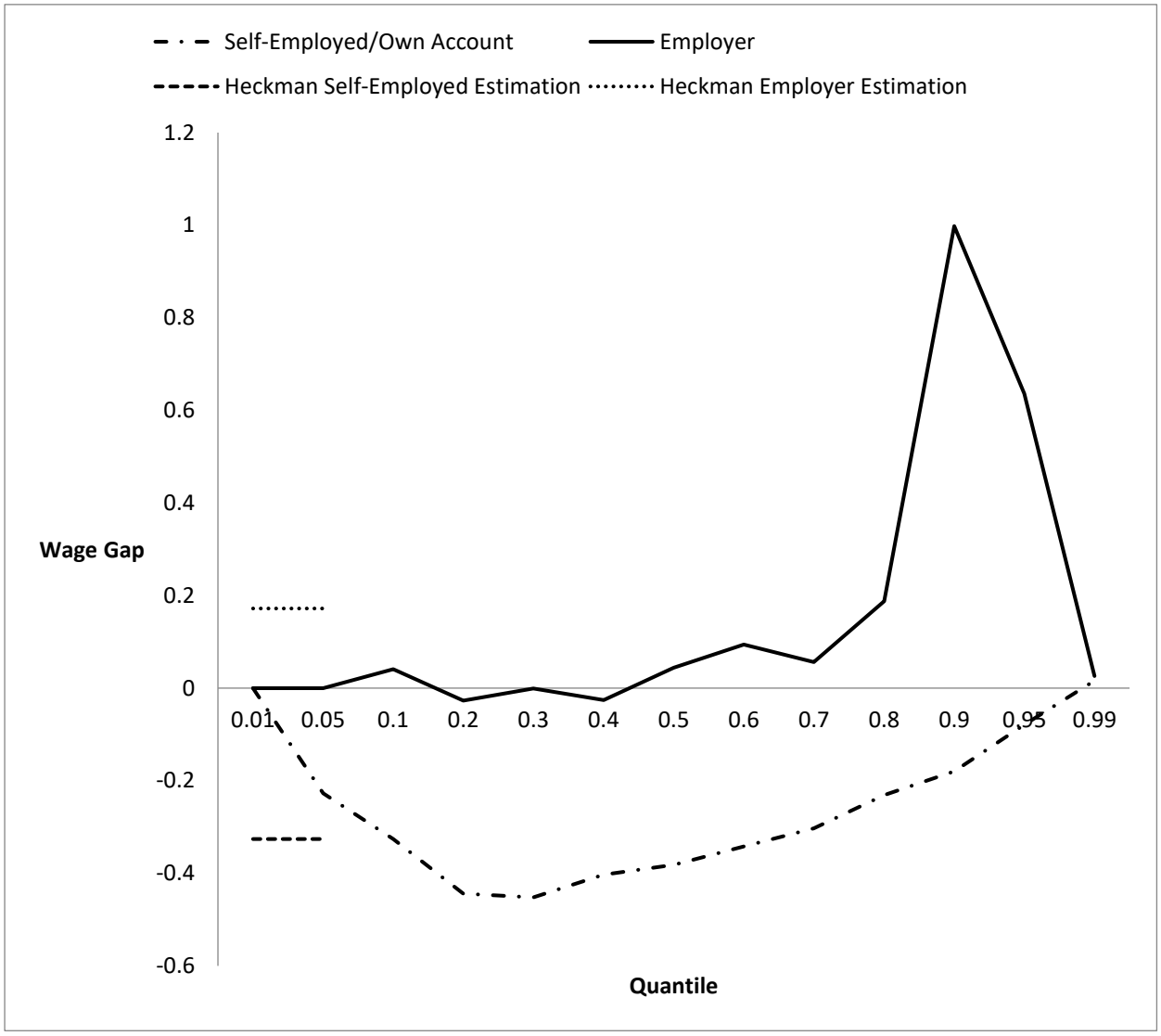


Figure 3. Heckman and Quantile Labour Income Premiums and Penalties – Female Sample



References

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: journal of the Econometric Society*, 153-161.

Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of economic perspectives*, 15(4), 143-156.