

**Misinformed, mismatched or magical?
Decomposing the gap between expected and realized
wages among graduates in Mozambique**

Sam Jones, Ricardo Santos, Gimelgo Xirinda

UNU-WIDER, Mozambique

05 November 2019

Agenda

- 1** Introduction
- 2** Framework
- 3** Background + Data
- 4** Results
- 5** Digging deeper
- 6** Summary

(1) Introduction

Motivation

- Systematically biased future expectations encountered in many settings
- Labour market: **expected wages** > **realized wages**
 - Weinstein (1980): +21.6% US college students (self vs other)
 - Smith & Powell (1990): +17% error among US undergrads
 - Avitabile & de Hoyos (2018): +33% error among Mexican high schoolers
- Pertinent since human capital investments made on basis of expected returns (Becker, 1964) :- erroneous expectations \implies resource misallocation
- Not so clear *why* positive bias ('unrealistic optimism') arises or persists
- We address this gap, using the structure of elicited expectations to identify proximate sources (types) of error
- Novel decomposition, using longitudinal data \implies **which types of errors matter**

Where might expectational errors come from?

In theory, 3 broad types of error:

1 Misinformation:

- About returns to specific jobs in labour market
- About returns to individual characteristics

2 Mismatch into labour market positions:

- *Vertical*: required vs actual education
- *Horizontal*: field of study vs field of work
- *Temporal*: time to complete studies
- Important since mismatches typically associated with material wage penalties (McGuinness et al., 2018; Somers et al., 2019)
... + pathways to 'good' jobs \neq clear.

3 Over-/under- confidence (systematic bias)

Previous studies have often documented the *presence* of aggregate expectational errors; but none have provided a more nuanced classification.

(2) Framework

Proximate determinants of earnings

Starting point: (subjective) own-wage expectations are almost always of a **conditional** form:

$$w_{ij}^e = E(w_{ij} \mid O^e, \Omega^e)$$

i.e., expectations are conditional on outcomes (the desired job) and perceived rewards to these same outcomes.

To put empirical structure on this, use a Mincerian (hedonic) wage function:

$$\begin{aligned} W_{ijt} &= e^{\mu + \delta t} Z_{it}^{\beta} H_{jt}^{\gamma} \epsilon_{it} \\ \ln W_{ijt} \equiv w_{ijt} &= \mu + \delta t + z_{it}\beta + h_{jt}\gamma + \epsilon_{it} \\ \implies w_{ij}^e &= \mu^e + \delta^e t_i^e + z_i^e \beta^e + h_j^e \gamma^e + \epsilon_{ij}^e \end{aligned}$$

So, this means we have:

$$\underbrace{\Omega^e = \{\mu^e, \delta^e, \beta^e, \gamma^e\}}_{\text{Expected rewards}}; \quad \underbrace{O^e = \{t_i^e, Z_i^e, H_j^e\}}_{\text{Expected outcomes}}$$

Expectational error decomposition

Comparing expected vs. realized wages gives the **expectational error**:

$$\underbrace{w_i^e - w_i^r}_{\text{Overall error}} = (\mu^e - \mu^r) + (t_i^e \delta^e - t_i^r \delta^r) + (z_i^e \beta^e - z_i^r \beta^r) + (h_j^e \gamma^e - h_j^r \gamma^r) + (\varepsilon_i^e - \varepsilon_i^r)$$

Noting that: $z_i^e \beta^e - z_i^r \beta^r = z_i^e \Delta \beta + \Delta z_i \beta^r$ (c.f., Blinder-Oaxaca)

Gives the error decomposition:

$$\ln W_i^e - \ln W_i^r \equiv \Delta w_{it} = e_i^I + e_i^M + [e_i^C + \Delta \varepsilon_{it}]$$

$$e_i^I = (t_i^e \Delta \delta + z_i^e \Delta \beta) + h_j^e \Delta \gamma \quad (2a)$$

$$e_i^M = \Delta t_i \delta^r + \Delta z_i \beta^r + \Delta H_j \gamma^r \quad (2b)$$

$$e_i^C = \Delta \mu \quad (2c)$$

Four specific types of error

- 1 $e_i^{I(j)}$: information regarding rewards to **job characteristics**
- 2 $e_i^{I(i)}$: information regarding rewards to **individual characteristics**
- 3 e_i^M : job **match quality** (outcomes)
- 4 e_i^C : **systematic bias** (c.f., in macro., **optimism** as shocks to TFP)

(3) Background + Data

Application to Mozambique

Relevant aspects of country context:

- Significant human capital deficit, reflecting legacy of colonialism and subsequent conflict
- Rapid growth of tertiary education over past decades (30% per year), from low base:
 - 700 new graduates in 2003 → 18,000 in 2016
- Challenging jobs environment:
 - 300,000 young people entering labour market each year
 - only 12% of all workers earn a wage
 - current real GDP growth barely matches population growth

Longitudinal survey

- Baseline survey in 2017 of final year undergraduates in 6 major universities in the country, public and private
- Sample representative by university, study area and gender
- Initial sample = 2,176 students, of which 1,989 provided valid wage expectations information
- 2018–2019, 6 waves of follow-up via mobile phone \implies we cover \sim 18 months post-study
- Low attrition: 88% reached in the last round and only 31 never contacted
- Focus here on value of *first wage* reported during post-study follow-up period vs. expected first wage reported at baseline

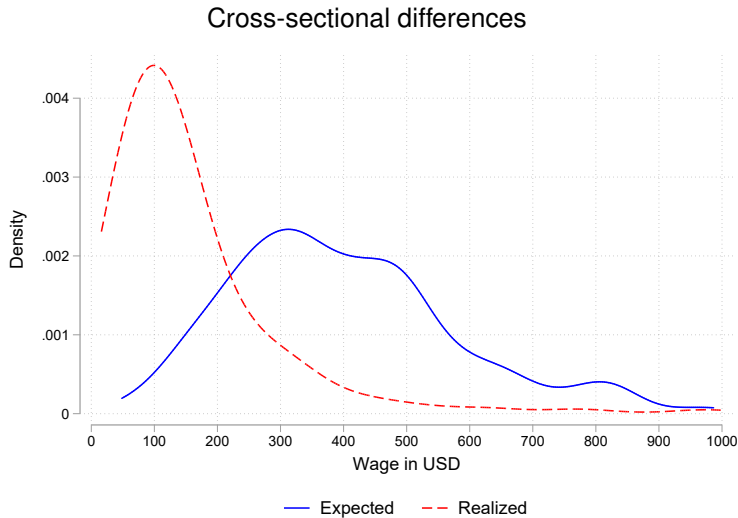
Baseline descriptive statistics

	Obtained work post-study?				All	
	No		Yes			
<i>Individual characteristics:</i>						
Age	24.61	(0.21)	25.41	(0.15)	25.31	(0.13)
Female	0.63	(0.02)	0.45	(0.01)	0.47	(0.01)
Married	0.10	(0.01)	0.12	(0.01)	0.12	(0.01)
Has kids	0.21	(0.02)	0.27	(0.01)	0.26	(0.01)
<i>University / course:</i>						
Public university	0.69	(0.02)	0.80	(0.01)	0.79	(0.01)
Total cost USD/month	78.05	(2.95)	63.98	(1.41)	65.71	(1.26)
Education	0.21	(0.02)	0.30	(0.01)	0.29	(0.01)
Humanities	0.01	(0.00)	0.02	(0.00)	0.02	(0.00)
Social Sciences	0.56	(0.02)	0.46	(0.01)	0.47	(0.01)
Natural Sciences	0.04	(0.01)	0.04	(0.01)	0.04	(0.00)
Engineering	0.06	(0.01)	0.07	(0.01)	0.07	(0.01)
Agriculture	0.06	(0.01)	0.05	(0.01)	0.05	(0.01)
Health	0.05	(0.01)	0.06	(0.01)	0.06	(0.01)
Observations	484		1408		1,892	

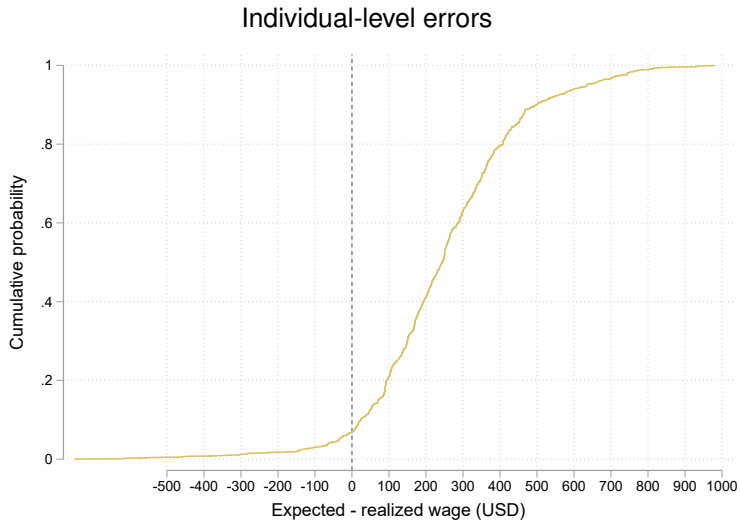
Realized outcomes in first paid position (N = 1,892)

	Private uni.		Public uni.		All
	Male	Female	Male	Female	
Private sector employee	0.65	0.72	0.55	0.59	0.59
Public employee	0.16	0.09	0.17	0.17	0.16
NGO employee	0.09	0.04	0.10	0.06	0.08
Self employed	0.08	0.11	0.16	0.15	0.15
Study unfinished	0.61	0.59	0.74	0.64	0.68
Job unlike course	0.61	0.65	0.66	0.70	0.67
Intern position	0.15	0.20	0.13	0.14	0.14
Works part time	0.47	0.38	0.62	0.51	0.55
No fixed contract	0.66	0.60	0.69	0.69	0.68
Searching for work	0.64	0.59	0.69	0.64	0.66
Employee mismatch	0.65	0.72	0.74	0.74	0.73
Sector mismatch	0.44	0.45	0.65	0.59	0.59
Mismatch count	3.91	3.99	4.53	4.27	4.33
Realized wage (USD/month)	222.30	201.07	145.73	134.67	154.68
Expected - realized wage (USD)	265.38	213.17	286.35	223.85	255.22
Expectational error (log.)	0.97	0.83	1.29	1.14	1.16

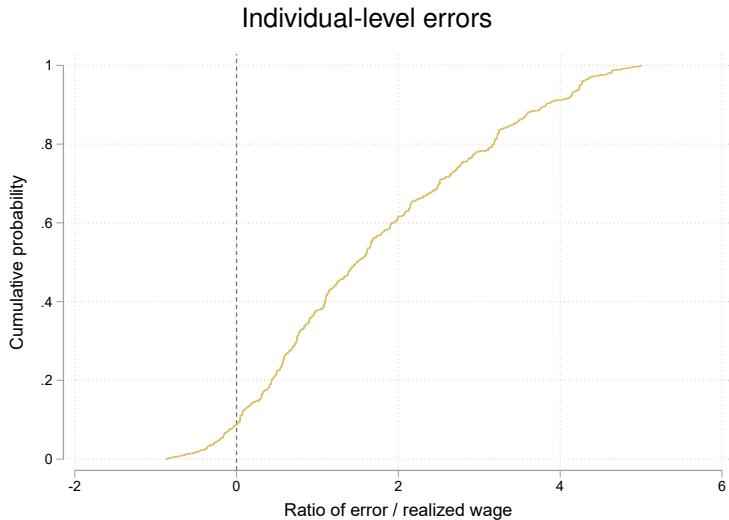
Expected vs. realized wages



Expected vs. realized wages



Expected vs. realized wages



(4) Results

Results

- Levels regression: ▶ Determinants of wages
- Error regression: ▶ Error decomposition
- Decomposition: ▶ Error components
- Figure 1: ▶ Mean error components
- Figure 2: ▶ Error component distributions
- Figure 3a: ▶ Subcomponents job chars. error
- Figure 3b: ▶ Subcomponents indiv chars. error
- Figure 3c: ▶ Subcomponents match quality error
- Figure 4: ▶ Errors by mismatch count
- Figure 5: ▶ Errors by quantile of expectational errors

	(I) Job?	(II) Expected wage	(III) Realized wage		
Constant	0.89*** (0.07)	3.15*** (0.18)	3.08*** (0.20)	1.65*** (0.17)	2.42*** (0.18)
Female	-0.08*** (0.02)	-0.16*** (0.03)	-0.07 (0.07)	0.10* (0.06)	-0.03 (0.08)
Private university	-0.06** (0.03)	0.05 (0.05)	0.10 (0.06)	0.34*** (0.08)	0.30*** (0.09)
Education	0.04* (0.02)	-0.02 (0.05)	-0.06 (0.06)	-0.14* (0.08)	-0.22*** (0.07)
Natural Sciences	-0.03 (0.03)	0.12** (0.06)	0.17*** (0.06)	0.10 (0.10)	0.13 (0.08)
Engineering	-0.03 (0.04)	0.18* (0.09)	0.18* (0.10)	0.23* (0.13)	0.33*** (0.10)
Health	0.03 (0.05)	0.34*** (0.07)	0.29*** (0.07)	0.29* (0.15)	0.09 (0.11)
English proficiency	0.04* (0.03)	-0.03 (0.05)	-0.09 (0.06)	0.14 (0.09)	0.19** (0.09)
Academic level (self)	0.03* (0.02)	0.01 (0.03)	-0.04 (0.04)	0.14*** (0.05)	0.08* (0.05)
Self employed	-0.00 (0.02)	-0.02 (0.04)	0.01 (0.05)	-0.06 (0.07)	-0.39*** (0.08)
Study unfinished					-0.25*** (0.07)
Works part time					-0.29*** (0.07)
Job unlike course					-0.20*** (0.06)
Obs.	1,891	1,891	1,401	1,401	1,401
R ²	0.07	0.15	0.16	0.20	0.33
Actual outcomes?	No	No	No	No	Yes

	(I) OLS		(II) Robust	
Constant	1.47*** (0.19)	1.04*** (0.25)	1.43*** (0.23)	0.95*** (0.25)
Female	-0.25*** (0.07)	-0.05 (0.11)	-0.25*** (0.06)	-0.12 (0.09)
Prev. work exp.	0.03*** (0.01)	0.02 (0.02)	0.04*** (0.01)	0.02* (0.01)
Private university	-0.31*** (0.08)	-0.21** (0.10)	-0.32*** (0.07)	-0.21** (0.09)
English proficiency	-0.19* (0.10)	-0.26** (0.11)	-0.17** (0.07)	-0.25*** (0.08)
Academic level (self)	-0.15*** (0.05)	-0.13** (0.05)	-0.12*** (0.05)	-0.08 (0.05)
Self employed	0.07 (0.09)	0.43*** (0.12)	0.05 (0.07)	0.36*** (0.09)
Study unfinished (Δ)		-0.17** (0.07)		-0.19*** (0.06)
Works part time (Δ)		-0.27*** (0.08)		-0.27*** (0.05)
Job unlike course (Δ)		-0.15*** (0.06)		-0.19*** (0.05)
NGO employee (Δ)		0.30*** (0.09)		0.31*** (0.09)
Self employed (Δ)		-0.39*** (0.09)		-0.34*** (0.07)
Nonselection hazard		-0.12 (0.09)		-0.06 (0.08)
Obs.	1,401	1,401	1,401	1,401
R ²	0.12	0.24	0.14	0.29

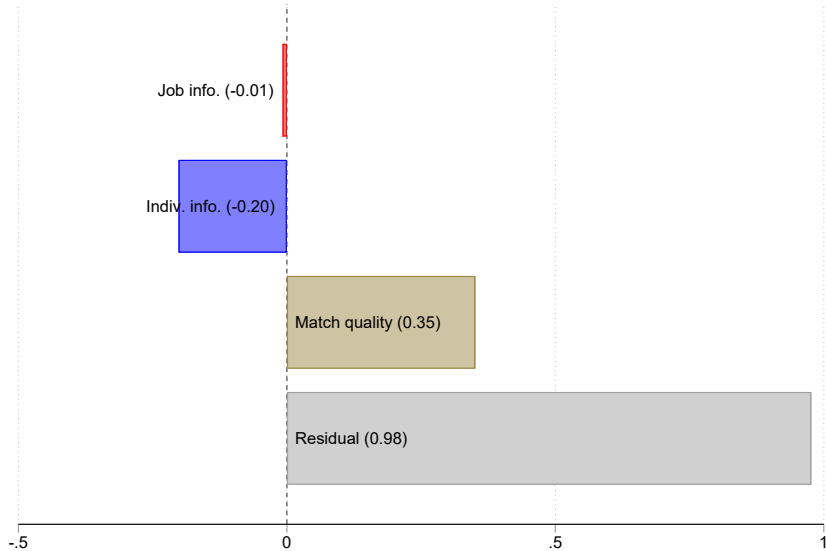
Error components

Combine terms, using a shrinkage approach – e.g.,:

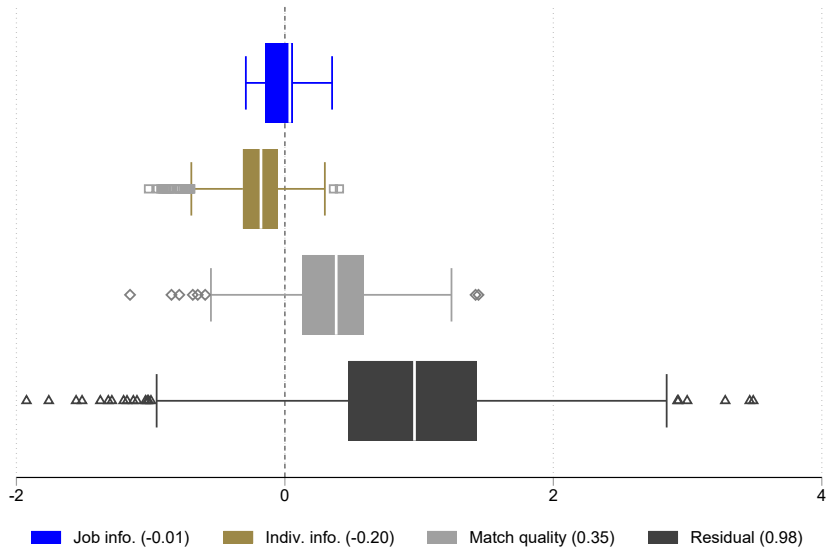
$$\hat{e}_i^M = \sum_{x \in \Delta t, \Delta Z, \Delta H} x_i \times \hat{\theta}_x \times [1 - \Pr(\hat{\theta}_x = 0)] \quad (3)$$

	(I) OLS		(II) Robust	
Job info.	0.18	0.02	0.11	-0.01
	[0.0,0.3]	[-0.2,0.2]	[0.0,0.2]	[-0.1,0.1]
Indiv. info.	-0.39	-0.15	-0.40	-0.20
	[-0.6,-0.2]	[-0.3,0.0]	[-0.5,-0.3]	[-0.4,-0.0]
Match quality	0.00	0.21	0.00	0.35
	[.,.]	[-0.1,0.5]	[.,.]	[0.2,0.5]
Syst. bias	1.40	1.01	1.42	0.98
	[1.0,1.8]	[0.4,1.6]	[1.0,1.9]	[0.5,1.5]

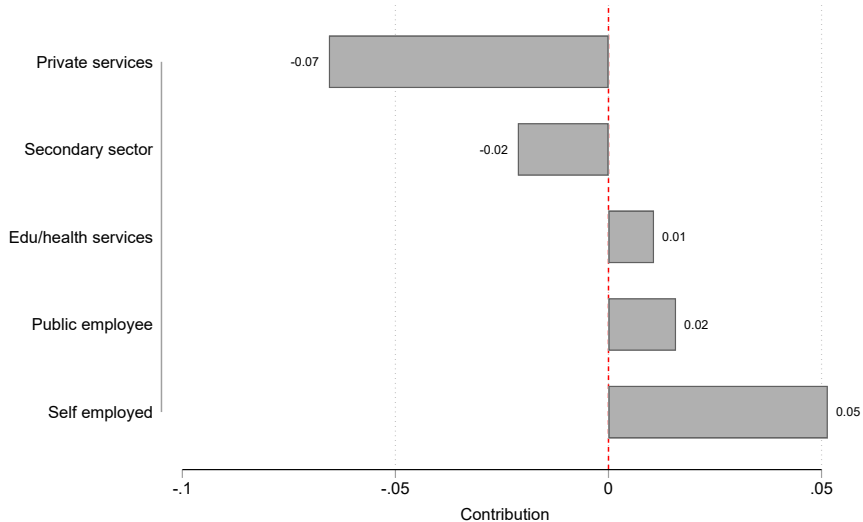
Mean error components



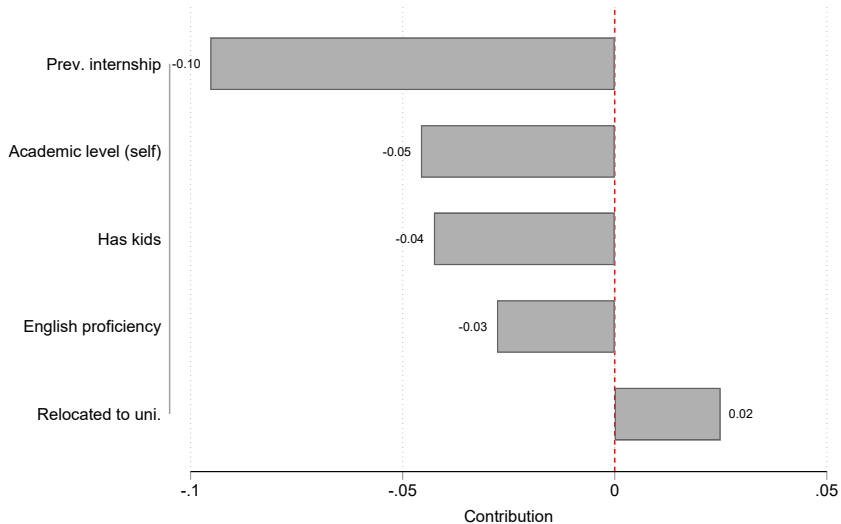
Error component distributions



Subcomponents of job info. error

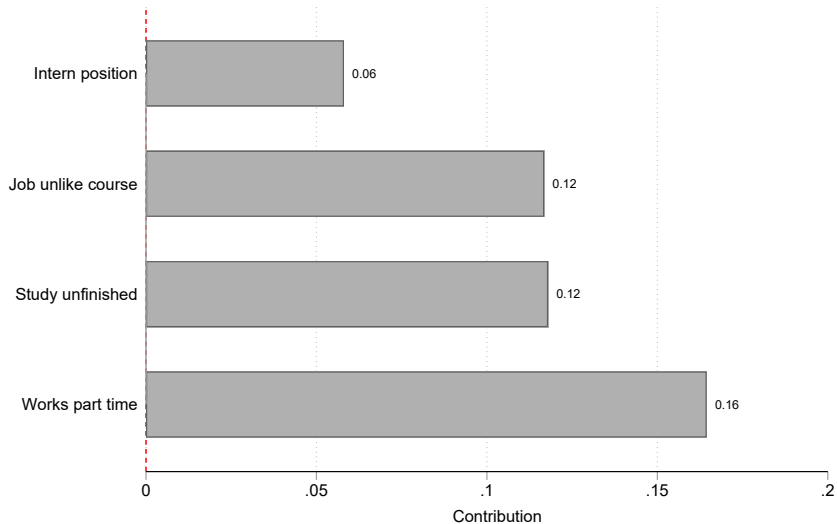


Subcomponents of individual info. error



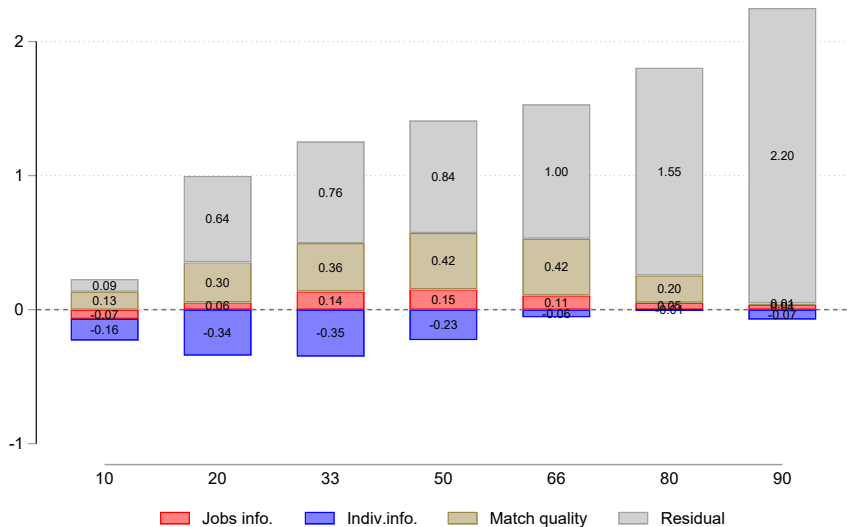
Total error = -0.20

Subcomponents of match quality error



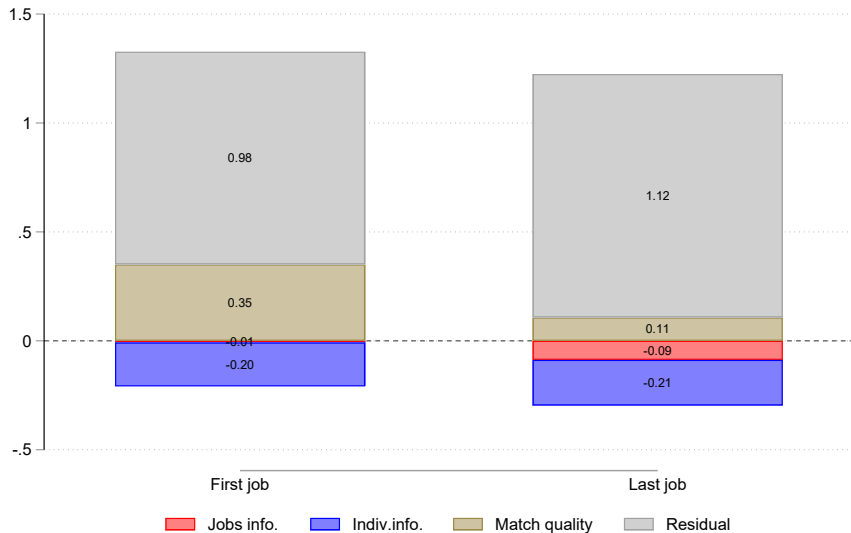
Total error = 0.35

Errors by quantile of expectational errors



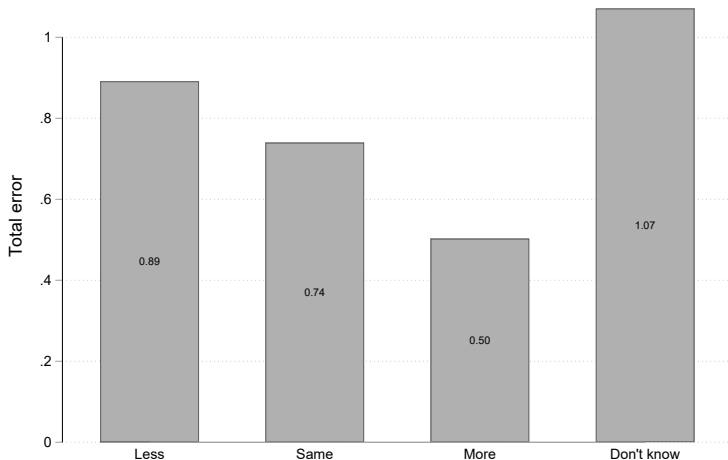
(5) Digging deeper

Do mismatches persist?



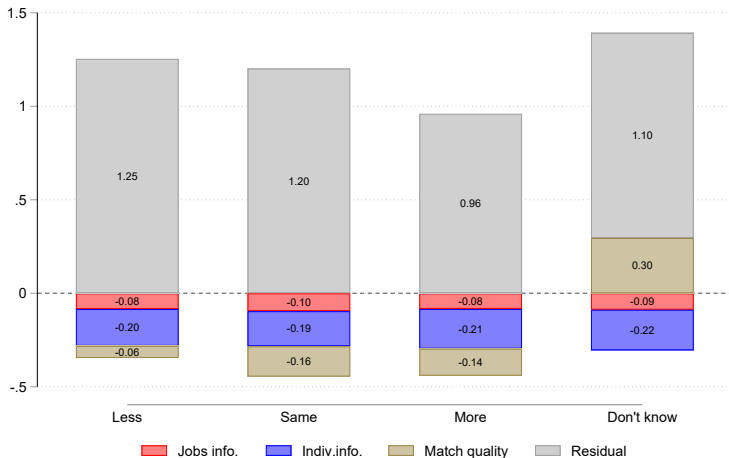
What *is* the systematic bias?

Is your current wage less, same or more than what you had expected when we asked you in 2017?



What *is* the systematic bias?

Even those earning more than they thought they had expected in 2017 show large positive residual component \Rightarrow original optimism was weakly-held / 'magical'.



(6) Summary

Summary

Contributions:

- 1 Go beyond aggregate errors to shed light on relevant types (sources) of error
- 2 Practical decomposition leveraging the conditional structure of expected wages
- 3 First longitudinal study of expectational errors among graduates in low income country (Mozambique)

Highlights:

- 1 Overall, expectational errors are very large ($> 100\%$)
- 2 Specific informational errors not so important, even negative w.r.t. indiv. chars
- 3 Errors due to job mismatch are large and prevalent, accounting for $\approx 33\%$ of expectational error in first wage in post-study period, but fall gradually
- 4 Systematic optimism is substantial, appears much larger than in many other contexts (although, comparable in magnitude to optimism of illegal immigrants)

Summary

Finally, some broader implications:

- 1 Key challenge is to further understand and (perhaps) address mismatches, which are indicative of significant market frictions & demand-side constraints
 - Students have some info. about labour market rewards ...
 - But less capacity to navigate opportunities and secure 'good' job posts
- 2 Systematic bias is unrealistic but not *necessarily* a cause for concern:
 - Appears fairly superficial / weak
 - Likely reflects a degree of magical thinking
- 3 Further work on *how* expectations are formed is necessary (i.e., are expectations updated based on new info.?)