Misinformed or mismatched?

Decomposing the gap between expected and realized wages among graduates in Mozambique

Sam Jones, Ricardo Santos, Gimelgo Xirinda

UNU-WIDER, Mozambique

11 September 2019

Agenda

(1) Introduction

Motivation

- Systematically biased future expectations encountered in many settings
- Labour market: expected wages > realized wages
- 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 $\Rightarrow$ **which types of errors matter**

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 $\Longrightarrow$ 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 $\Rightarrow$ **which types of errors matter**

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 $\Rightarrow$ **which types of errors matter**

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 $\Rightarrow$ **which types of errors matter**

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 $\Rightarrow$ **which types of errors matter**

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 $\Rightarrow$ **which types of errors matter**

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 $\Rightarrow$ **which types of errors matter**

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 $\Rightarrow$ **which types of errors matter**

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 $\Rightarrow$ **which types of errors matter**

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. Mismatch into labour market positions:

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. Mismatch into labour market positions:

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. Mismatch into labour market positions:

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. 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.

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. 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.

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, 'self-enhancement' bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. 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.

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. 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.

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. 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.

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. 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.

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. 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.

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

Where might expectational errors come from?

In theory, 4 main types of error:

1. Over-confidence, ‘self-enhancement’ bias
2. Incomplete information regarding returns in labour market
3. Incomplete information regarding returns to individual characteristics
4. 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.

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:

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

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:

$$W_{ijt} = e^{\mu+\delta t}\ Z_{it}^{\beta}\ H_{jt}^{\gamma}\ \epsilon_{it}$$

$$\ln W_{ijt} = w_{ijt} = \mu + \delta t + z_{it}\beta + h_{jt}\gamma + \varepsilon_{it}$$

$$\implies w_{ij}^{e} = \mu^{e} + \delta^{e} t_{i}^{e} + z_{i}^{e}\beta^{e} + h_{j}^{e}\gamma^{e} + \varepsilon_{ij}^{e}$$

So, this means we have:

$$\underbrace{\Omega^{e} = \{\mu^{e}, \delta^{e},\ \beta^{e}, \gamma^{e}\}}_{\text{Expected rewards}};\quad \underbrace{O^{e} = \{t_{i}^{e}, Z_{i}^{e}, H_{j}^{e}\}}_{\text{Expected outcomes}}$$

Proximate determinants of earnings

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

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

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:

$$W_{ijt} = e^{\mu+\delta t}\ Z_{it}^{\beta}\ H_{jt}^{\gamma}\ \epsilon_{it}$$

$$\ln W_{ijt} = w_{ijt} = \mu + \delta t + z_{it}\beta + h_{jt}\gamma + \varepsilon_{it}$$

$$\implies w_{ij}^e = \mu^e + \delta^e t_i^e + z_i^e\beta^e + h_j^e\gamma^e + \varepsilon_{ij}^e$$

So, this means we have:

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

Proximate determinants of earnings

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

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

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:

$$W_{ijt} = \mathrm{e}^{\mu+\delta t}\ Z_{it}^{\beta}\ H_{jt}^{\gamma}\ \epsilon_{it}$$

$$\ln W_{ijt} \equiv w_{ijt} = \mu + \delta t + z_{it}\beta + h_{jt}\gamma + \varepsilon_{it}$$

$$\implies w_{ij}^e = \mu^e + \delta^e t_i^e + z_i^e \beta^e + h_j^e \gamma^e + \varepsilon_{ij}^e$$

So, this means we have:

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

Proximate determinants of earnings

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

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

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:

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

So, this means we have:

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

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:

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

$$e_i^P = \Delta\mu \tag{2a}$$

$$e_i^I = (t_i^e \Delta\delta + z_i^e \Delta\beta) + h_j^e \Delta\gamma \tag{2b}$$

$$e_i^M = \Delta t_i \delta^r + \Delta z_i \beta^r + \Delta H_j \gamma^r \tag{2c}$$

Four sources / types of error

1. e_i^P: **generic optimism** (c.f., macro., optimism as shocks to TFP)

2. $e_i^{I(j)}$: information regarding rewards to **job characteristics**

3. $e_i^{I(i)}$: information regarding rewards to **individual characteristics**

4. e_i^M: job **match quality** (outcomes)

Four sources / types of error

1. e_i^P: **generic optimism** (c.f., macro., optimism as shocks to TFP)

2. $e_i^{I(j)}$: information regarding rewards to **job characteristics**

3. $e_i^{I(i)}$: information regarding rewards to **individual characteristics**

4. e_i^M: job **match quality** (outcomes)

Four sources / types of error

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

Four sources / types of error

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

(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 $\rightarrow$ 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

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

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 $\rightarrow$ 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

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 $\rightarrow$ 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

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 $\rightarrow$ 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

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 $\rightarrow$ 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

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 $\rightarrow$ 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, 4 waves of follow-up via mobile phone (2 further waves planned) $\Longrightarrow$ here we cover 12 months post-study
- Low attrition: 1,887 followed-up at least one (5.1% lost/refused)
- Focus here on value of *first wage* reported during post-study follow-up period vs. expected first wage reported at baseline

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, 4 waves of follow-up via mobile phone (2 further waves planned) $\Longrightarrow$ here we cover 12 months post-study
- Low attrition: 1,887 followed-up at least one (5.1% lost/refused)
- Focus here on value of *first wage* reported during post-study follow-up period vs. expected first wage reported at baseline

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, 4 waves of follow-up via mobile phone (2 further waves planned) $\Longrightarrow$ here we cover 12 months post-study
- Low attrition: 1,887 followed-up at least one (5.1% lost/refused)
- Focus here on value of *first wage* reported during post-study follow-up period vs. expected first wage reported at baseline

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, 4 waves of follow-up via mobile phone (2 further waves planned) $\Longrightarrow$ here we cover 12 months post-study
- Low attrition: 1,887 followed-up at least one (5.1% lost/refused)
- Focus here on value of *first wage* reported during post-study follow-up period vs. expected first wage reported at baseline

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, 4 waves of follow-up via mobile phone (2 further waves planned) $\Longrightarrow$ here we cover 12 months post-study
- Low attrition: 1,887 followed-up at least one (5.1% lost/refused)
- Focus here on value of *first wage* reported during post-study follow-up period vs. expected first wage reported at baseline

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, 4 waves of follow-up via mobile phone (2 further waves planned) $\Longrightarrow$ here we cover 12 months post-study
- Low attrition: 1,887 followed-up at least one (5.1% lost/refused)
- Focus here on value of *first wage* reported during post-study follow-up period vs. expected first wage reported at baseline

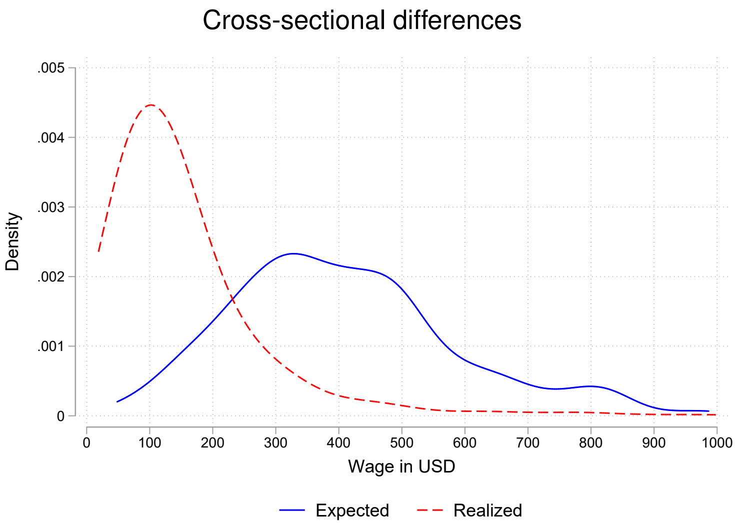
Baseline descriptive statistics

	Obtained work post-study?					
	No		Yes		All	
Individual characteristics:						
Age	24.42	(0.17)	26.93	(0.20)	26.05	(0.14)
Female	0.60	(0.02)	0.36	(0.01)	0.44	(0.01)
Married	0.09	(0.01)	0.18	(0.01)	0.14	(0.01)
Has kids	0.20	(0.02)	0.37	(0.01)	0.31	(0.01)
University / course:						
Public university	0.71	(0.02)	0.85	(0.01)	0.80	(0.01)
Total cost USD/month	73.68	(2.34)	62.34	(1.49)	66.31	(1.28)
Education	0.24	(0.02)	0.36	(0.01)	0.32	(0.01)
Humanities	0.01	(0.00)	0.02	(0.00)	0.02	(0.00)
Social Sciences	0.51	(0.02)	0.40	(0.01)	0.44	(0.01)
Natural Sciences	0.04	(0.01)	0.04	(0.01)	0.04	(0.00)
Engineering	0.07	(0.01)	0.08	(0.01)	0.07	(0.01)
Agriculture	0.05	(0.01)	0.06	(0.01)	0.05	(0.01)
Health	0.07	(0.01)	0.06	(0.01)	0.06	(0.01)
Observations	700		1,187		1,887	

Realized outcomes in first paid position (N = 1,887)

	Private uni.		Public uni.		
	Male	Female	Male	Female	All
Private sector employee	0.57	0.62	0.42	0.46	0.46
Public employee	0.21	0.12	0.27	0.33	0.27
NGO employee	0.07	0.04	0.09	0.05	0.07
Self employed	0.11	0.16	0.19	0.14	0.17
Study unfinished	0.79	0.78	0.86	0.82	0.83
Job unlike course	0.55	0.63	0.50	0.57	0.54
Intern position	0.13	0.18	0.11	0.11	0.12
Works part time	0.43	0.38	0.48	0.38	0.44
No fixed contract	0.73	0.66	0.74	0.71	0.72
Searching for work	0.69	0.63	0.68	0.58	0.65
Employee mismatch	0.62	0.67	0.68	0.60	0.65
Sector mismatch	0.41	0.39	0.52	0.44	0.48
Mismatch count	3.88	4.02	4.09	3.79	3.98
Realized wage (USD/month)	226.23	196.23	149.85	139.17	156.21
Expected - realized wage (USD)	255.31	228.51	293.60	239.67	270.37
Expectational error (log.)	0.94	0.92	1.27	1.13	1.18

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	0.71***	3.17***	3.17***	1.65***	2.36***
	(0.09)	(0.15)	(0.17)	(0.16)	(0.17)
Female	-0.18***	-0.15***	-0.07	0.07	0.02
	(0.03)	(0.03)	(0.06)	(0.06)	(0.09)
Private university	-0.14***	0.04	0.09	0.35***	0.34***
	(0.03)	(0.05)	(0.08)	(0.09)	(0.11)
Education	0.02	-0.03	-0.02	-0.12**	-0.20***
	(0.03)	(0.05)	(0.05)	(0.06)	(0.06)
Engineering	0.02	0.22**	0.19	0.29	0.32**
	(0.06)	(0.09)	(0.12)	(0.19)	(0.13)
Academic level (self)	0.05*	0.03	-0.01	0.14***	0.07*
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Public employee	-0.01	-0.05*	-0.04	-0.23***	-0.06
	(0.03)	(0.03)	(0.04)	(0.05)	(0.07)
Self employed	0.01	-0.02	0.04	-0.04	-0.32***
	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)
Nonselection hazard			-0.11*		-0.02
			(0.06)		(0.08)
Study unfinished					-0.28***
					(0.07)
Works part time					-0.32***
					(0.06)
Job unlike course					-0.17***
					(0.05)
Obs.	1,887	1,887	1,187	1,187	1,187
R^2	0.20	0.14	0.15	0.23	0.34
Actual outcomes?	No	No	No	No	Yes

	(I) OLS		(II) Robust	
Constant	1.52***	0.80***	1.61***	0.88***
	(0.18)	(0.24)	(0.23)	(0.26)
Female	-0.24***	-0.09	-0.21***	-0.07
	(0.08)	(0.10)	(0.06)	(0.10)
Prev. work exp.	0.03***	0.02	0.03***	0.02
	(0.01)	(0.02)	(0.01)	(0.02)
Private university	-0.35***	-0.23**	-0.34***	-0.21**
	(0.07)	(0.10)	(0.08)	(0.10)
Health	0.33***	0.32***	0.35***	0.35***
	(0.11)	(0.08)	(0.13)	(0.12)
Academic level (self)	-0.12***	-0.09*	-0.12***	-0.08*
	(0.05)	(0.04)	(0.05)	(0.05)
Self employed	0.09	0.33***	0.06	0.24***
	(0.08)	(0.10)	(0.07)	(0.09)
Study unfinished (Δ)		-0.23***		-0.24***
		(0.07)		(0.07)
Works part time (Δ)		-0.35***		-0.37***
		(0.07)		(0.07)
Job unlike course (Δ)		-0.13***		-0.18***
		(0.05)		(0.05)
NGO employee (Δ)		0.21**		0.28***
		(0.09)		(0.09)
Self employed (Δ)		-0.29***		-0.22***
		(0.07)		(0.07)
Obs.	1,187	1,187	1,187	1,187
R^2	0.14	0.24	0.15	0.28

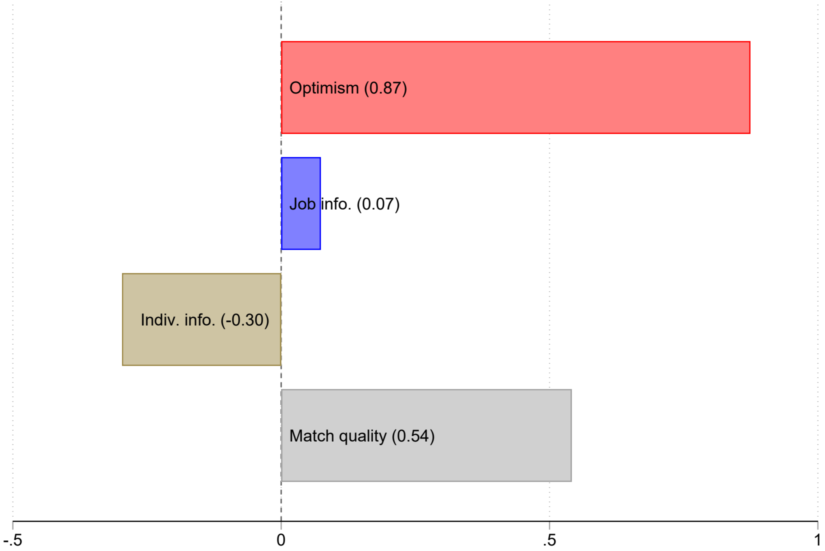
Error components

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

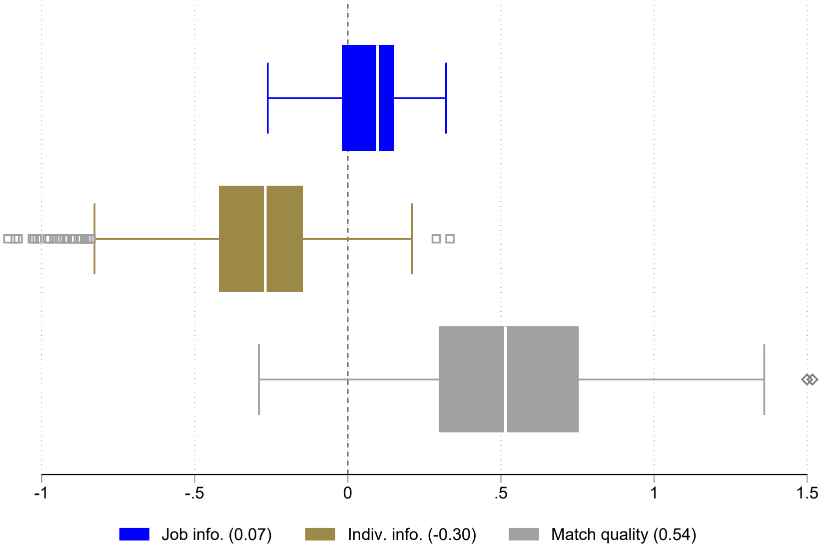
$$\hat{e}_i^M = \sum_{x \in \Delta t, \Delta Z, \Delta H} x_i \times \hat{\theta}_x \times [1 - \Pr(\hat{\theta}_x = 0)] \tag{3}$$

	(I) OLS		(II) Robust	
Optimism	1.52	0.80	1.61	0.87
	[1.2,1.9]	[0.3,1.3]	[1.2,2.1]	[0.4,1.4]
Job info.	0.11	0.12	0.08	0.07
	[0.0,0.2]	[-0.0,0.3]	[0.0,0.1]	[-0.1,0.2]
Indiv. info.	-0.40	-0.25	-0.42	-0.30
	[-0.5,-0.2]	[-0.4,-0.1]	[-0.6,-0.3]	[-0.4,-0.2]
Match quality	0.00	0.51	0.00	0.54
	[.,.]	[0.4,0.7]	[.,.]	[0.4,0.7]

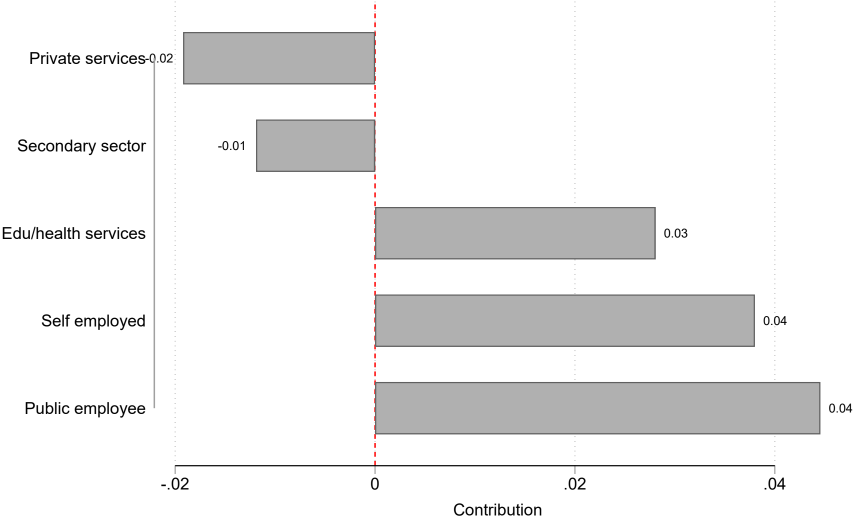
Mean error components



Error component distributions

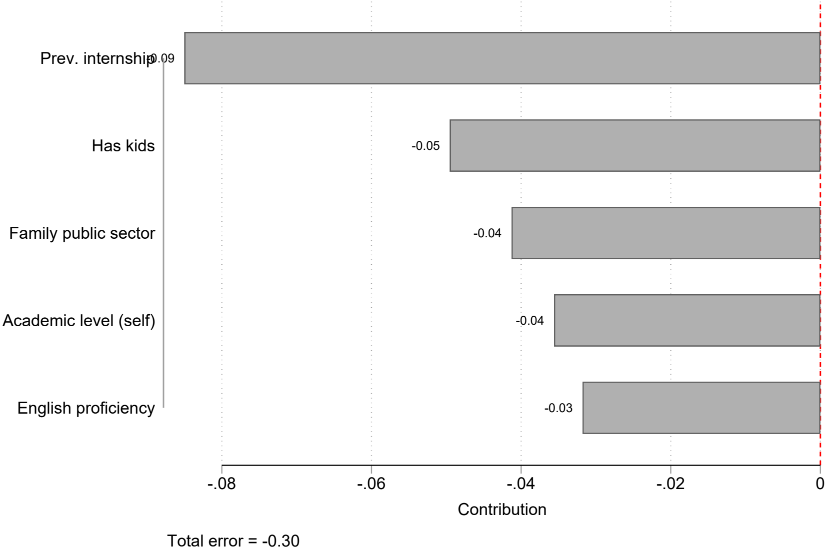


Subcomponents of job info. error

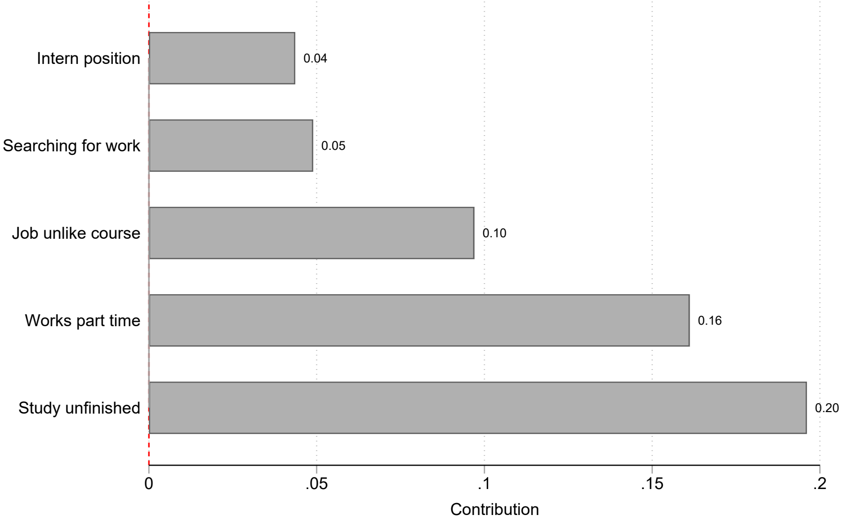


Total error = 0.07

Subcomponents of individual info. error

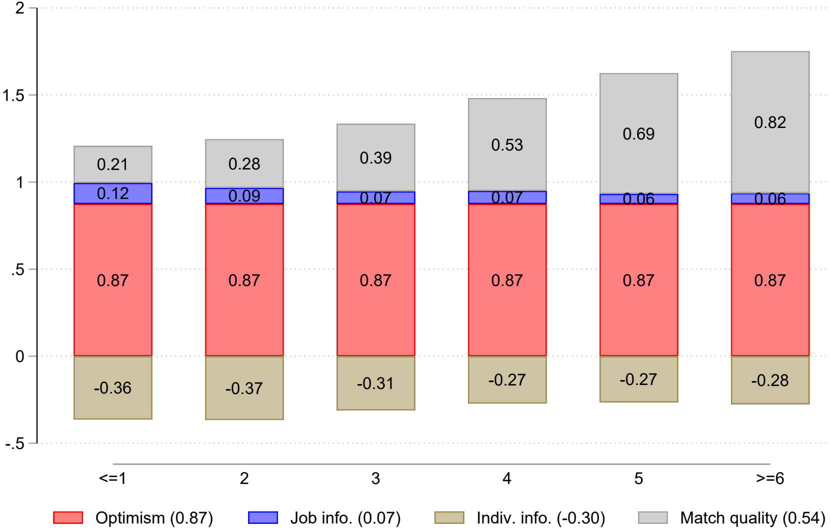


Subcomponents of match quality error

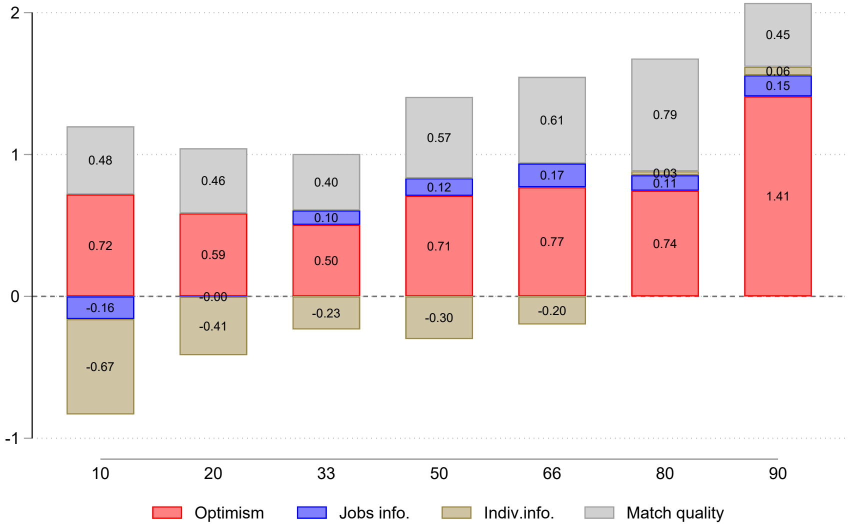


Total error = 0.54

Errors by mismatch count



Errors by quantile of expectational errors



(5) 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 $\approx 50\%$ of expectational error in first wage in post-study period
4. Generic optimism (productivity) is also substantial, much larger than elsewhere

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 $\approx 50\%$ of expectational error in first wage in post-study period
4. Generic optimism (productivity) is also substantial, much larger than elsewhere

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 $\approx 50\%$ of expectational error in first wage in post-study period
4. Generic optimism (productivity) is also substantial, much larger than elsewhere

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 $\approx 50\%$ of expectational error in first wage in post-study period
4. Generic optimism (productivity) is also substantial, much larger than elsewhere

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 $\approx 50\%$ of expectational error in first wage in post-study period
4. Generic optimism (productivity) is also substantial, much larger than elsewhere

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 $\approx 50\%$ of expectational error in first wage in post-study period
4. Generic optimism (productivity) is also substantial, much larger than elsewhere

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 $\approx 50\%$ of expectational error in first wage in post-study period
4. Generic optimism (productivity) is also substantial, much larger than elsewhere

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. Magnitude of generic optimism **may be** a cause for concern (e.g., potential source of youth frustration), but difficult to interpret *per se*

   – Does not appear to be *only* self-enhancement bias

   – Perhaps reflects continuation of economic crisis (in part)

3. Future work on how expectations are formed is necessary (i.e., are expectations updated based on new info.?)

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. Magnitude of generic optimism **may be** a cause for concern (e.g., potential source of youth frustration), but difficult to interpret *per se*

   – Does not appear to be *only* self-enhancement bias

   – Perhaps reflects continuation of economic crisis (in part)

3. Future work on how expectations are formed is necessary (i.e., are expectations updated based on new info.?)

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. Magnitude of generic optimism **may be** a cause for concern (e.g., potential source of youth frustration), but difficult to interpret *per se*
   - Does not appear to be *only* self-enhancement bias
   - Perhaps reflects continuation of economic crisis (in part)
3. Future work on how expectations are formed is necessary (i.e., are expectations updated based on new info.?)

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. Magnitude of generic optimism **may be** a cause for concern (e.g., potential source of youth frustration), but difficult to interpret *per se*

   – Does not appear to be *only* self-enhancement bias

   – Perhaps reflects continuation of economic crisis (in part)

3. Future work on how expectations are formed is necessary (i.e., are expectations updated based on new info.?)

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. Magnitude of generic optimism **may be** a cause for concern (e.g., potential source of youth frustration), but difficult to interpret *per se*
   - Does not appear to be *only* self-enhancement bias
   - Perhaps reflects continuation of economic crisis (in part)
3. Future work on how expectations are formed is necessary (i.e., are expectations updated based on new info.?)

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. Magnitude of generic optimism **may be** a cause for concern (e.g., potential source of youth frustration), but difficult to interpret *per se*
   - Does not appear to be *only* self-enhancement bias
   - Perhaps reflects continuation of economic crisis (in part)
3. Future work on how expectations are formed is necessary (i.e., are expectations updated based on new info.?)

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. Magnitude of generic optimism **may be** a cause for concern (e.g., potential source of youth frustration), but difficult to interpret *per se*
   - Does not appear to be *only* self-enhancement bias
   - Perhaps reflects continuation of economic crisis (in part)
3. Future work on how expectations are formed is necessary (i.e., are expectations updated based on new info.?)