Learning management through matching: A field experiment using mechanism design

Girum Abebe (World Bank), Marcel Fafchamps (Stanford), Michael Koelle (OECD), Simon Quinn (Oxford)
A novel field experiment

Can aspiring entrepreneurs acquire management skills by observing managers in large firms?
A novel field experiment

Can aspiring entrepreneurs acquire management skills by observing managers in large firms?

We placed aspiring Ethiopian entrepreneurs into established medium and large firms:

- We used random assignment to participate of firms and individuals.
- We assigned individuals to host firms with a Gale-Shapley matching algorithm.
Can aspiring entrepreneurs acquire management skills by observing managers in large firms?

We placed aspiring Ethiopian entrepreneurs into established medium and large firms:

- We used random assignment to participate of firms and individuals.
- We assigned individuals to host firms with a Gale-Shapley matching algorithm.

Our research design enables us to estimate two different types of treatment effects:
Can aspiring entrepreneurs acquire management skills by observing managers in large firms?

We placed aspiring Ethiopian entrepreneurs into established medium and large firms:

- We used random assignment to participate of firms and individuals.
- We assigned individuals to host firms with a Gale-Shapley matching algorithm.

Our research design enables us to estimate two different types of treatment effects:

1. ATE / ITT of the assignment to participating in the programme;
Can aspiring entrepreneurs acquire management skills by observing managers in large firms?

We placed aspiring Ethiopian entrepreneurs into established medium and large firms:

- We used random assignment to participate of firms and individuals.
- We assigned individuals to host firms with a Gale-Shapley matching algorithm.

Our research design enables us to estimate two different types of treatment effects:

1. ATE / ITT of the assignment to participating in the programme;
2. Heterogeneous effects based on the performance of the matching algorithm.
Contribution to literature

Heterogeneity in management: individual managers’ traits and experiences (Bertrand and Schoar, 2003; Ellison and Holden, 2013; Bemelech and Frydman, 2015; Kaplan, Klebanov and Sorensen, 2013; Bandiera, Hansen, Prat and Sadun, 2017) and management practices at the level of the organization (Bloom and van Reenen, 2007).

• We implement and analyse an intervention that changes individuals’ managerial capital; organizational management practices are an important mediator.

Applying theoretical insights and approaches from mechanism design to field experiments in developing countries (Jayachandran, de Laat, Lambin, Stanton, Audy and Thomas, 2017; Rigol, Hussam and Roth, 2018).

• We utilize mechanism design in the service of causal inference, similar to the school choice literature (Abdulkadiroğlu, Angrist, Narita and Pathak, 2017)

• We show how mechanism design can improve program effectiveness over ad-hoc matching methods (Trapp, Teytelboym, Martinello, Anderson and Ahani, 2018)
The programme

We invited **young Ethiopians** (aged 18 to 30, inclusive), having a minimum of technical/vocational, college or university qualifications.
The programme

We invited young Ethiopians (aged 18 to 30, inclusive), having a minimum of technical/vocational, college or university qualifications.

We advertised through social media, campus visits, and ‘job boards’, using the following headline message:

Do you want to be your own boss?
See how successful firms work!
Gain business and management skills first hand!
Context: Young labour force entrants

Our sample consists of young, highly educated and highly motivated Ethiopians shortly after graduating from tertiary education:

- 75% male, and 75% have a college degree (most frequently, engineering or business).
Context: Young labour force entrants

Our sample consists of young, highly educated and highly motivated Ethiopians shortly after graduating from tertiary education:

- 75% male, and 75% have a college degree (most frequently, engineering or business).
- 50% graduated in year before placement, or in the same year.
Context: Young labour force entrants

Our sample consists of **young, highly educated** and **highly motivated** Ethiopians shortly after graduating from tertiary education:

- **75% male**, and **75% have a college degree** (most frequently, engineering or business).
- **50% graduated in year before** placement, or in the same year.
- **80% actively search** for a wage job, and **30% plan to start or expand a business**.

<table>
<thead>
<tr>
<th>Wave</th>
<th>Self-employed</th>
<th>Wage employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>(t = 0)</td>
<td>7%</td>
</tr>
<tr>
<td>6 months</td>
<td>(t = 1)</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>(t = 2)</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Control group.*
Context: Young labour force entrants

Our sample consists of **young, highly educated** and **highly motivated** Ethiopians shortly after graduating from tertiary education:

- **75% male**, and **75% have a college degree** (most frequently, engineering or business).
- **50% graduated in year before** placement, or in the same year.
- **80% actively search** for a wage job, and **30% plan to start or expand a business**.

<table>
<thead>
<tr>
<th>Wave</th>
<th>Self-employed</th>
<th>Wage employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>((t = 0))</td>
<td>7%</td>
</tr>
<tr>
<td>6 months</td>
<td>((t = 1))</td>
<td>10%</td>
</tr>
<tr>
<td>12 months</td>
<td>((t = 2))</td>
<td>13%</td>
</tr>
</tbody>
</table>

*Note: Control group.*
Context: Young labour force entrants

Our sample consists of **young, highly educated** and **highly motivated** Ethiopians shortly after graduating from tertiary education:

- **75% male**, and **75% have a college degree** (most frequently, engineering or business).
- **50% graduated in year before** placement, or in the same year.
- **80% actively search** for a wage job, and **30% plan to start or expand a business**.

<table>
<thead>
<tr>
<th>Wave</th>
<th>Self-employed</th>
<th>Wage employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline ($t = 0$)</td>
<td>7%</td>
<td>25%</td>
</tr>
<tr>
<td>6 months ($t = 1$)</td>
<td>10%</td>
<td>59%</td>
</tr>
<tr>
<td>12 months ($t = 2$)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Control group.*
Context: Young labour force entrants

Our sample consists of young, highly educated and highly motivated Ethiopians shortly after graduating from tertiary education:

- **75% male**, and **75% have a college degree** (most frequently, engineering or business).
- **50% graduated in year before** placement, or in the same year.
- **80% actively search** for a wage job, and **30% plan to start or expand a business**.

<table>
<thead>
<tr>
<th>Wave</th>
<th>Self-employed</th>
<th>Wage employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline ($t = 0$)</td>
<td>7%</td>
<td>25%</td>
</tr>
<tr>
<td>6 months ($t = 1$)</td>
<td>10%</td>
<td>59%</td>
</tr>
<tr>
<td>12 months ($t = 2$)</td>
<td>13%</td>
<td>69%</td>
</tr>
</tbody>
</table>

*Note: Control group.*
The Experiment

Management experience placements (‘internships’):

- **four weeks** in a medium to large firm, mostly in Addis Ababa;
- required **full-time**, daily commitment at the firm;
- paid a small stipend (25th percentile of baseline wages).
The Experiment

Management experience placements (‘internships’):

- **four weeks** in a medium to large firm, mostly in Addis Ababa;
- required **full-time**, daily commitment at the firm;
- paid a small stipend (25th percentile of baseline wages).

We used **pairwise randomisation**, stratified on gender, education and age.
Firms were randomised too.

- Treated firms **hosted 1-5 interns** (median and mode: 2).
- Firms operate in **services** (about 40%), **manufacturing** (about 25%), **trade** (about 20%) and other sectors.
- The median firm has **57 employees** \(Q_1 = 22; Q_3 = 155\).
Matching interns and firms

Our experiment features about **1650 applicants**, of whom about **825** were assigned to internships. These interns were hosted by about **350 firms**.
Matching interns and firms

Our experiment features about 1650 applicants, of whom about 825 were assigned to internships. These interns were hosted by about 350 firms.

For logistical reasons, we implemented on a rolling basis, using a total of 42 batches (i.e. an average of about 20 interns per batch).
Matching interns and firms

Our experiment features about **1650 applicants**, of whom about **825** were assigned to internships. These interns were hosted by about **350 firms**.

For logistical reasons, we implemented on a **rolling basis**, using a total of **42 batches** (i.e. an average of about 20 interns per batch).

Within each batch, we ask all **interns to rank all firms**, and all **firms to rank all interns**. We then match interns and firms with a **deferred-acceptance stable matching algorithm**, in which firms propose (Gale and Shapley, 1962).

- Firms rank interns based on a short, anonymous CV;
- Interns rank firms based on: name, sector, location, size.
Empirical specification

We collect follow-up data using face-to-face surveys at six months and twelve months after treatment.

Our preferred estimating equation is ANCOVA with pairwise dummies; that is, for individual $i$ in pair $p$ at time $t > 0$, we estimate:

$$y_{ipt} = \beta_1 \cdot T_i + \beta_2 \cdot y_{ip0} + \delta_p + \epsilon_{ipt}.$$  \hspace{1cm} (1)

We conduct inference on the ITT coefficient $\beta_1$ as following:

- We cluster at the individual level.
- We report Wald $p$-values, and false-discovery rate $q$-values (Benjamini, Krieger and Yekutieli, 2006) within families of outcomes.

We filed a pre-analysis plan at www.socialscienceregistry.org/trials/2776.
Primary outcome: **Occupation**

<table>
<thead>
<tr>
<th>Dummy: Treated</th>
<th>(1) Self-employed</th>
<th>(2) Self-emp. hours</th>
<th>(3) Profit income</th>
<th>(4) Wage work</th>
<th>(5) Perm. work</th>
<th>(6) Managerial work</th>
<th>(7) Wage work hours</th>
<th>(8) Wage income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.00 (0.01) [0.72] [0.45]</td>
<td>-0.01 (0.08) [0.87] [0.48]</td>
<td>108.44 (180.66) [0.55] [0.38]</td>
<td>0.03 (0.02) [0.05] * [0.07] *</td>
<td>0.04 (0.02) [0.01] ** [0.03] **</td>
<td>0.02 (0.01) [0.15]</td>
<td>0.41 (0.14) [0.00] ***</td>
<td>265.25 (88.50) [0.00] ***</td>
</tr>
<tr>
<td>Control mean (follow-up)</td>
<td>0.12</td>
<td>0.71</td>
<td>923.07</td>
<td>0.64</td>
<td>0.51</td>
<td>0.12</td>
<td>4.91</td>
<td>2520.80</td>
</tr>
<tr>
<td>Control mean (baseline)</td>
<td>0.07</td>
<td>0.35</td>
<td>438.47</td>
<td>0.26</td>
<td>0.19</td>
<td>0.04</td>
<td>1.76</td>
<td>853.33</td>
</tr>
<tr>
<td>Observations</td>
<td>3,110</td>
<td>3,121</td>
<td>3,077</td>
<td>3,121</td>
<td>3,121</td>
<td>3,121</td>
<td>3,121</td>
<td>3,105</td>
</tr>
</tbody>
</table>
Primary outcome: **Management**

<table>
<thead>
<tr>
<th>Family:</th>
<th>(1) Confidence</th>
<th>(2)</th>
<th>(3) Management Practices</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome:</td>
<td>Sum</td>
<td>Index</td>
<td>Overall</td>
<td>Marketing</td>
<td>Recording</td>
<td>Financial</td>
</tr>
<tr>
<td>Dummy: Treated</td>
<td>0.23</td>
<td>0.04</td>
<td>0.08</td>
<td>0.07</td>
<td>0.11</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.01)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.09)</td>
<td>(0.06)</td>
</tr>
<tr>
<td></td>
<td>[0.00]*****</td>
<td>[0.00]*****</td>
<td>[0.09]*</td>
<td>[0.22]</td>
<td>[0.19]</td>
<td>[0.47]</td>
</tr>
<tr>
<td></td>
<td>{0.00}***</td>
<td>{0.00}***</td>
<td>{0.42}</td>
<td>{0.42}</td>
<td>{0.42}</td>
<td>{0.42}</td>
</tr>
<tr>
<td>Control mean (follow-up)</td>
<td>9.78</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.00</td>
<td>-0.05</td>
<td>-0.01</td>
</tr>
<tr>
<td>Control mean (baseline)</td>
<td>9.61</td>
<td>-0.04</td>
<td>0.07</td>
<td>0.02</td>
<td>0.17</td>
<td>0.02</td>
</tr>
<tr>
<td>Observations</td>
<td>3,121</td>
<td>3,121</td>
<td>396</td>
<td>396</td>
<td>396</td>
<td>396</td>
</tr>
</tbody>
</table>

**Note:** For the management practices outcome family, we run OLS and omit the pairwise dummies for randomization strata.
We test whether hosting an intern...

1. ... changed firms’ stated preferences about future interns;

2. ... changed firms’ management practices and labour flows;

3. ... caused interns to have more similar attitudes to their hosts.
We test whether hosting an intern...

1. ...changed firms’ stated preferences about future interns;

2. ...changed firms’ management practices and labour flows;

3. ...caused interns to have more similar attitudes to their hosts.

We find **no effect** on any of these outcomes.
We used a **controlled mechanism** to assign interns to firms.

We find a **positive effect of being treated**.
We used a **controlled mechanism** to assign interns to firms.

We find a **positive effect of being treated**.

Does the effect depend upon the mechanism? Could an alternative mechanism have generated larger effects?
We used a controlled mechanism to assign interns to firms.

We find a positive effect of being treated.

Does the effect depend upon the mechanism? Could an alternative mechanism have generated larger effects?

To fix ideas...

What is the occupational effect of being assigned to a ‘high-management’ host rather than a ‘low-management’ host?
Identifying the causal effect of assignment to treatment varieties

In our mechanism, we **control**:

1. The assignment of both sides of the market (‘firms’ and ‘interns’) given rankings;
2. The information set either side has to rank the other side;
3. The grouping of interns into small batches, based on calling applicants from a locally randomly ordered list.
Identifying the causal effect of assignment to treatment varieties

In our mechanism, we **control**:

1. The assignment of both sides of the market (‘firms’ and ‘interns’) given rankings;
2. The information set either side has to rank the other side;
3. The grouping of interns into small batches, based on calling applicants from a locally randomly ordered list.

When a placement is assigned using a ‘fair’ centralised assignment mechanism with an oversubscription lottery, conditioning on the **propensity score eliminates selection bias** and the setting becomes equivalent to a stratified randomised experiment (Abdulkadiroğlu, Angrist, Narita and Pathak, 2017).
Identifying the causal effect of assignment to treatment varieties

In our mechanism, we **control:**

1. The assignment of both sides of the market (‘firms’ and ‘interns’) given rankings;
2. The information set either side has to rank the other side;
3. The grouping of interns into small batches, based on calling applicants from a locally randomly ordered list.

When a placement is assigned using a ‘fair’ centralised assignment mechanism with an oversubscription lottery, conditioning on the **propensity score eliminates selection bias** and the setting becomes equivalent to a stratified randomised experiment (Abdulkadiroğlu, Angrist, Narita and Pathak, 2017).

**The idea:** We can obtain the propensity score of a **deterministic mechanism** (Gale-Shapley Deferred Acceptance) by treating the composition of other interns as a **random variable**, integrated out by simulation.
How did our mechanism assign interns to firms?

Firm $f$’s ranking over interns $I$:

$$r_{fI} \sim \rho_f(w_1, w_2, \ldots);$$

$w_j$: characteristics of intern $j$;
How did our mechanism assign interns to firms?

Firm $f$’s ranking over interns $I$:

$$r_{fI} \sim \rho_f(w_1, w_2, \ldots);$$

$w_j$: characteristics of intern $j$;

Intern $i$’s ranking over firms $F$:

$$r_{iF} \sim \tau_i(x_1, x_2, \ldots);$$

$x_k$: characteristics of firm $k$;
How did our mechanism assign interns to firms?

Firm $f$’s ranking over interns $I$:

$$r_{fI} \sim \rho_f(w_1, w_2, \ldots);$$

$w_j$: characteristics of intern $j$;

Intern $i$’s ranking over firms $F$:

$$r_{iF} \sim \tau_i(x_1, x_2, \ldots);$$

$x_k$: characteristics of firm $k$;

Assignments $m_{IF}$ are determined by mechanism $\psi$:

$$m_{IF} = \psi(R_{IF}, R_{FI}),$$

$$[r_{iF}, R_{-i,F}]$$

*Notation:* we stack into bold $\rho_F$ the functionals for the set of firms $F$, etc.
The $i$-conditional propensity score

$$p_{if} \equiv \Pr(m_{if} = 1 \mid r_{iF}, w_i, X_F, \rho_f)$$

This is the \textbf{propensity score} of the assignment to a firm, conditional on $Z_{iF} = \{r_{iF}, w_i, X_F, \rho_F\}$
The $i$-conditional propensity score

$$p_{if} \equiv \Pr(m_{if} = 1 \mid r_{iF}, w_i, X_F, \rho_f)$$

$$= \int \psi_{if} \left( [r_{iF}, \tau_{-i}(X_F)] \right) \rho_F \left( [w_i, W_{-i}] \right) \, dF(W_{-i}, \tau_{-i}).$$

This is the **propensity score** of the assignment to a firm, conditional on $Z_{iF} = \{r_{iF}, w_i, X_F, \rho_F\}$ which consists of:

1. intern $i$’s observed ranking $r_{iF}$ and observables $w_i$ we give to the firms;
2. all firms $F$’s preferences $\rho_F$ and observables $X_F$ we give to interns.
The $i$-conditional propensity score

$$p_{if} \equiv \Pr(m_{if} = 1 \mid r_{iF}, w_i, X_F, \rho_f)$$

$$= \int \psi_{if} \left( \left[ r_{iF}, \tau_{-i}(X_F) \right], \rho_F \left( [w_i, W_{-i}] \right) \right) \, dF(W_{-i}, \tau_{-i}).$$

This is the **propensity score** of the assignment to a firm, conditional on $Z_{iF} = \{r_{iF}, w_i, X_F, \rho_F\}$ which consists of:

1. intern $i$’s observed *ranking* $r_{iF}$ and observables $w_i$ we give to the firms;
2. all firms $F$’s preferences $\rho_F$ and observables $X_F$ we give to interns.

We integrate over other potential interns’ $-i$ preferences and characteristics.
The \textit{i}-conditional propensity score

\[ p_{if} \equiv \Pr(m_{if} = 1 \mid r_{iF}, w_i, X_F, \rho_f) \]

\[ = \int \psi_{if}\left([r_{iF}, \tau_{-i}(X_F)]\right), \rho_F\left([w_i, W_{-i}]\right) dF(W_{-i}, \tau_{-i}) \]

This is the \textbf{propensity score} of the assignment to a firm, conditional on
\[ Z_{iF} = \{r_{iF}, w_i, X_F, \rho_F\} \] which consists of:

1. intern \textit{i}'s observed \textit{ranking} \( r_{iF} \) and observables \( w_i \) we give to the firms;
2. all firms \( F \)'s preferences \( \rho_F \) and observables \( X_F \) we give to interns.

We integrate over \textit{other potential interns'} \( -i \) preferences and characteristics.

We can sum across \( p_{if} \) to obtain the propensity score for assignment of \( i \) to a particular \textit{type} of firm, \( D_i \).
A Bayesian approach to simulating the propensity score

We need to integrate over the joint distribution of other interns’ characteristics and preferences: $F(w_j, \tau_j) \forall j \notin i$. This is an object we do not directly observe and need to estimate.
A Bayesian approach to simulating the propensity score

We need to integrate over the joint distribution of other interns’ characteristics and preferences: \( F(w_j, \tau_j) \forall j \notin i \). This is an object we do not directly observe and need to estimate.

Our estimation combines:

1. A random coefficient random utility model (rank-ordered logit) with a finite support of types empirically replaces the unknown functionals \( \rho_f \) and \( \tau_i \).
2. A weak Dirichlet prior on the marginal distribution of intern observables \( w_j \).
A Bayesian approach to simulating the propensity score

We need to integrate over the joint distribution of other interns’ characteristics and preferences: $F(w_j, \tau_j) \forall j \not\in i$. This is an object we do not directly observe and need to estimate.

Our estimation combines:

1. A random coefficient random utility model (rank-ordered logit) with a finite support of types empirically replaces the unknown functionals $\rho_f$ and $\tau_i$.
2. A weak Dirichlet prior on the marginal distribution of intern observables $w_j$.

We estimate this generative statistical model in a Bayesian way, using MCMC estimation.

We then draw repeatedly from the corresponding posterior distributions of parameters, form rankings, and create assignments using the DA mechanism.
Marginal Treatment Effects under matching: **Profit earnings**
Marginal Treatment Effects under matching: Self-employment
Marginal Treatment Effects under matching: Wage employment
Total earnings under firm-proposing DA

\[ Y(DA-F) = \int \left[ p \cdot y_1(p) + (1 - p) \cdot y_0(p) \right] f(p) \, dp \]
What if we had used a **different mechanism**?

Denote the mechanism that we actually used — Deferred Acceptance with firms proposing — as ‘**Mechanism A**’.

Suppose that we are considering using some other mechanism: ‘**Mechanism B**’.
What if we had used a different mechanism?

Denote the mechanism that we actually used — Deferred Acceptance with firms proposing — as ‘Mechanism A’.

Suppose that we are considering using some other mechanism: ‘Mechanism B’.

Then we can repeat the previous integration, replacing the actual mechanism with the alternative mechanism. Then we can obtain:

\[
Y(\text{alternative}) = \int [p_b \cdot y_1(p_a, p_b) + (1 - p_b) \cdot y_0(p_a, p_b)] f(p_a, p_b) \, dp_a \, dp_b.
\]
What if we had used a different mechanism?

Denote the mechanism that we actually used — Deferred Acceptance with firms proposing — as ‘Mechanism A’.

Suppose that we are considering using some other mechanism: ‘Mechanism B’.

Then we can repeat the previous integration, replacing the actual mechanism with the alternative mechanism. Then we can obtain:

\[
Y(\text{alternative}) = \int [p_b \cdot y_1(p_a, p_b) + (1 - p_b) \cdot y_0(p_a, p_b)] f(p_a, p_b) \, dp_a \, dp_b.
\]

This relates to the Marginal Treatment Effect (Carneiro, Heckman and Vytlacil, 2011), and to the result that any treatment mean can be expressed as a weighted average of the Marginal Treatment Effect (Heckman and Vytlacil, 1999, 2005, 2007).
Bivariate distributions of propensity scores
Bivariate distributions of propensity scores
Bivariate distributions of propensity scores
Counter-factual mechanism results
Conclusion

We implemented and evaluated a novel field experiment matching individuals with firms to gain ‘management experience’.

• We find, on average, an increase in professional wage employment but not in planned or realized self-employment.
• We find some higher professed and demonstrated management ability.
We implemented and evaluated a novel field experiment matching individuals with firms to gain ‘management experience’.

- We find, on average, an increase in professional wage employment but not in planned or realized self-employment.
- We find some higher professed and demonstrated management ability.

We develop an empirical strategy for identifying how differences in host firms matter for interns.

- We find heterogeneous effects by host firm for self-employment, but not for wage employment.
- We find that the assignment mechanism matters profoundly for the average gain from the intervention.
- This methodology can be used as a starting point for field experiments involving heterogeneous treatment.
Introduction
Experiment & Context
Results: ATEs
Framework: Treatment varieties
Results: Variety MTEs
Other mechanisms