

なぜ男女差別は起こるのか？——差別の動学理論と実証分析

J. Aislinn Bohren, Alex Imas and Michael Rosenberg (2019) “The Dynamics of Discrimination: Theory and Evidence,” *American Economic Review*, 109 (10), 3395–3436.

東京大学大学院博士課程 鳥谷部貴大

性別や人種による差別は労働経済学における重要なテーマのひとつである。労働市場では、差別によって女性やマイノリティの人的資本蓄積の機会や十分に資質のある職業への就業機会が失われてしまう。こうして資源配分が歪められることによる経済的損失は無視できず、例えばアメリカの1960年以降の経済成長には女性や黒人が高技能な職業に就けるようになったことが大きく寄与している (Hsieh et al. 2019)。それではなぜ差別は行われるのだろうか？理論分析は多く行われており、雇用者による代表的な差別として選好にもとづく差別と信念にもとづく差別が挙げられる。実証分析では差別の存在自体は示されているものの (Altonji and Blank 1999)、差別のメカニズムを明らかにするような実証分析はあまり行われてこなかった。本稿で紹介する論文は差別のダイナミクスに着目することでメカニズムを特定する分析デザインを考案し、それを数学のQ&Aサイトを用いたフィールド実験によって検証している。

差別の動学モデル

ここでは労働者が每期タスクを行い、それが他者によって評価されるモデルを考える。労働者は性別と能力によって特徴付けられ、タスクの質は労働者の能力によって確率的に決まる。評価者は每期異なり、労働者の能力やタスクの質を観察できないが、労働者の性別、評価履歴、タスクの質に関するシグナルを観察できる。評価者はこれらをもとに労働者の能力に対する信念を更新して、タスクの質に自分の選好を加えたものの平均二乗誤差が最小になるように評価を行う。つまり、評価者の目的関数は $\min_v \hat{E}[(v - (q - c_g))^2 | h, s, g]$ である。ただし、 q 、 c_g 、 h 、 s 、 g はタスクの質、性別に対する選好、労働者の評価の履歴、シグナル、性別で、期待値は評価者の信念の上でとられる。

「差別」は同一のシグナルと評価履歴を持つ男女が受ける評価の差として定義され、ここでの差別の源泉は選好と信念である。また、評価者の信念は必ずしも正しいとは限らないことに注意が必要である。

このモデルから3つの予測が導出される。1) タスクの質を直接観察できる場合、選好にもとづく差別のみが起こる。2) すべての評価者が同じ女性差別的な信念を持っている場合、女性は常に差別される¹⁾。3) 「男性の方が能力が高い」という信念の評価者と「男女に能力差はない」という信念の評価者がいる場合²⁾、1期目に女性は差別されるが、2期目では十分に高い評価履歴を持つ労働者の間では男性が差別される。最初のふたつの予測は直感的だが、最後の予測はやや直感に反するかもしれない。予測3のケースでは、労働者の過去の評価は女性差別的な評価者によってなされた可能性があるため、差別的でない評価者は評価履歴が同じ男女なら女性の方が能力が高いと考え、女性を高く評価する。一方、差別的な評価者は2期目も女性を差別するものの、差別的でない評価者の割合が十分に大きければ、女性有利に働く効果の方が強く、高い評価履歴を持つ男女ではむしろ男性が差別されるのである。異なる信念を持つ評価者がいるということは、少なくとも1つの信念は誤りであり、これは誤った信念による差別であるといえる。

フィールド実験

これらの理論予測を検証するために、数学のQ&Aサイト Mathematics Stack Exchange を用いたフィールド実験が行われた。このサイトではユーザー同士で数学に関する質問や回答を投稿しあい、それぞれの投稿は他のユーザーによって評価される³⁾。回答の評価基準は「回答の正しさ」という客観的なものだが、質問については「その質問が自明でないか？興味深い

か？」という主観的な基準になっている。理論モデルと照らし合わせると、回答は質を観察できるタスク、質問は質を観察できないタスクだと考えられる。評価を行う画面には投稿者のユーザー名と評価数が表示されるため、ユーザー名と性別を結び付けることで評価履歴と差別の関係を分析できる。

質問への評価に関する実験を行うために、280の新規アカウントを作成し、140のアカウントに半分ずつ男性と女性のユーザー名を付ける。残りの140アカウントは評価数が上位25%になるまで活動した後、男女で半分ずつになるようにユーザー名を変更する。評価履歴と差別の関係を分析するために新規アカウントと高評価のアカウントを準備しているのである。そして、各アカウントに質問をランダムに割り振って投稿し、その評価数を記録する。次に、回答への評価に関する実験のために140の新規アカウントを作成して男女半分ずつになるようにユーザー名を付ける。そして、質問の検索と回答の作成を行い、ランダムに選ばれたアカウントから回答を投稿する。

まず、客観的な評価基準の設けられている回答に対する評価をみてみると、男女差はほとんどみられなかった。タスクの質を観察できる場合、選好にもとづく差別だけが問題になるので（理論予測1）、この結果はユーザーに差別選好がないことを示唆している。次に、質問に対する評価をみてみると、新規アカウントによる投稿の場合、男性に対する評価は女性よりも約0.4標準偏差が高かった。ユーザーに差別選好はないことをふまえると、この結果は、男性の方が女性よりも能力が高いと考えているユーザーがいることを意味している。では、すべてのユーザーがそのような考えを持っているのだろうか？もしそうであれば、すでに高い評価を得ているユーザーからの投稿についても同様の差別傾向が観察されるはずである（理論予測2）。しかし、高い評価を得ているアカウントによる投稿の場合、女性に対する評価は男性よりも約0.6標準偏差が高く、むしろ男性の方が差別されていることがわかった。この結果は理論予測3と整合的である。つまり、女性の能力は男性よりも低いと考えている評価者と女性の能力は男性と同等だと考えている評価者の両方がいることを示唆している。正しい信念はひとつなので、これは女性の能力に関する誤った信念による差別だといえる。

実験結果の解釈の懸念として、男女でサイトへの投

稿をやめる確率が異なる可能性が挙げられる。もしサイトへの投稿をやめる確率が性別と能力に依存するならば、新規アカウントの間では男性の能力の方が高かったとしても、継続的に投稿を続けているアカウントの間では女性の能力の方が高い可能性がある。すると、評価者が正しい信念を持っていても能力分布の変化に応じて差別傾向も変化するため、必ずしも実験結果を誤った信念にもとづく差別とは解釈できない。しかし、サイトから得られた観察データを用いた分析によって、投稿をやめる確率に男女差はみられないことがわかったため、実験結果は能力分布の変化によるものではないと考えられる。

むすびに

本論文は、数学のQ&Aサイトでのフィールド実験を通じて、男女差別は選好ではなく能力に関する誤った信念にもとづいていることを示しており、もし能力に性差がなければ、情報提供を通じて信念を更新させることで差別を緩和できる可能性がある。本論文から得られた知見をただちに労働市場における差別へ適用できるわけではないが、これまでなかなか明らかにされてこなかった差別のメカニズムを明らかにした点は重要である。効果的な差別は正政策を考えるうえで、差別のメカニズムを知ることは肝要であり、今後さらなる研究の蓄積が期待される。

- 1) 単純化のために、予測2と予測3では、評価者に差別的な選好はないケースを考えている。
- 2) 前者はすべての評価者の信念が自分と同じだと考えており、後者は評価者の信念の真の分布を知っていると仮定されている。
- 3) ユーザーは獲得した累積評価数に応じてサイト管理に関する権限が与えられる。また、評価を懸賞金として用いて質問を投稿することで、質の高い回答を集めることもできる。

参考文献

- Altonji, Joseph G. and Rebecca M. Blank (1999) "Chapter 48 Race and Gender in the Labor Market," *Handbook of Labor Economics*, Vol. 3, Elsevier, pp. 3143-3259.
- Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones and Peter J. Klenow (2019) "The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 87 (5), 1439-1474.

とりやべ・たかひろ 東京大学大学院経済学研究科博士課程。最近の主な論文に“Parental Leaves and Female Skill Utilization: Evidence from PIAAC,” *RIETI Discussion Paper*, 18-E-003 (2018年、共著)。労働経済学、応用ミクロ計量経済学専攻。