Candidate number: 49045

TITLE HERE

Supervisor: Ekaterina (Katya) Oparina
Word count:

Submitted as partial fulfilment for
MSc in Behavioural Science
Department of Psychological and Behavioural Science
The London School of Economics and Political Science

Total word limit: 10000

Abstract (100-150)

1 Introduction (1000)

Gender gap in labour force participation is a world-wide phenomenon which is particularly pronounced in developing countries. Globally, the rate of labour force participation is about 75% among men while only 50% among women **(International Labour Organization, 2021)**. In regions such as Middle East, North Africa, and South Asia, the gender gap is even greater, with around 75% of men and 20% of women participating in the labour force **(International Labour Organization, 2021)**. In the search of potential measures to narrow the gender gap, its causes have been extensively studied with a recent attention on the prominent role of social norms. Researchers speculate that interventions aiming at changing social norms may be the key to achieve greater gender equality in labour force participation **(Bursztyn et al., 2020; Codazzi et al., 2018; Jayachandran, 2021)**.

A recent successful attempt in developing such social norm interventions was made by Bursztyn and colleagues **(2020)** in Saudi Arabia. In Saudi Arabia, gender norms exist that expect women to be absent from labour or segregated from men at workplace, and that women need approval from their male "guardian" (usually the husband or father) if they want to work outside the home **(Bursztyn et al., 2018)**. These allow us to make a reasonable speculation that the low female labour participation rate **(28% in 2022, International Labour Organization)** may be because men don't allow their wives to work outside the home (WWOH) due to the internalisation of the gender norms (i.e., personal beliefs aligning with the norm that women should not work outside the home). Interestingly, however, Bursztyn and colleagues **(2020)** found that the critical factor was rather Saudi men's misperception of the injunctive norms regarding WWOH. Their findings revealed that 80-85% of surveyed Saudi men who were married and aged 18-35 reported to agree with the statement that women should be allowed to work outside the home, while more than three quarters of them underestimated this percentage.

Based on this result, Bursztyn and colleagues **(2020)** designed an intervention to correct men's misperception regarding WWOH found evidence supporting evidence for its effectiveness in changing their labour supply decisions. They found that men who received correct information of the true percentage of supporting men (vs. those who did not receive the information) were significantly more likely to sign up for job matching service for their wives immediately after the intervention. Their wives were also more likely to have applied and interviewed for a job three to five months after the intervention. The researchers also tested the effect of a similar intervention on Saudi women in a field setting, finding that women who received the information on the percentage of men supporting WWOH (vs. those who did not received the information) were more likely to take up a part-time job outside the home instead of a position to work at home.

The present research takes correcting norm misperception as a promising type of social norm intervention and seeks to explore its effectiveness when scaled up. Since large-scale policy intervention can usually be costly to implement and test, the present research aims to study, prior

to implementation, the intervention strategy that is theoretically the most effective under appropriate assumptions via agent-based modelling (ABM). The reason for choosing ABM as the method of modelling the system of WWOH action is that the dynamic interaction among people's private belief and norm perception regarding WWOH, as well as their labour supply decision and action can be viewed as complex **(Johnson, 2009)**. One's labour supply decision and action can presumably influence their acquaintances' private belief and norm perception, and these influences in turn impact their acquaintances' labour supply decision and action as a function of private belief and norm perception, eventually creating a feedback loop among individuals. This nature of having agents interacting dynamically and irregularly over time makes the system of WWOH action irreducible to lower-level descriptions (e.g., mathematical equations) and qualifies it as a complex system **(Gustafsson & Sternad, 2010)**. ABMs are suitable for studying the dynamics and emergent properties of such systems, as well as simulate and explore the consequences of policy interventions to them **(Bailey et al., 2019; Madsen et al., 2019)**.

give space for specific research questions

2 Literature Review (2000)

2.1 Models of Belief and Action Dynamics

Belief and action changes among a group of interconnected and mutually influencing agents have been studied widely through mathematical and computational models. These models assume that agents hold beliefs regarding certain issues. At each time step, these beliefs are modified based on their neighbours' beliefs according to some rule. The dynamics of belief distribution in a population over time is studied. This section will first review various methods to formally model agents' beliefs and actions and the ways in which agents influence one another. Then, the topics that these models are usually applied to studied are introduced.

2.1.1 Formal Representation of Beliefs and Actions

Different models of belief and action dynamics model agents' beliefs and actions and the ways in which they are affected by those of other agents differently **(Hassani et al., 2022)**. The most simplistic models represent agents' beliefs using binary variables, the classic ones of which include the Ising model **(Li et al., 2019)**, the voter model **(Holley & Liggett, 1975)**, and the Sznajd model **(Sznajd-Weron & Sznajd, 2000)**. Although these models and their variations adopt different rules to update agents' beliefs, the updating usually results in the agents being memoryless about their previous beliefs. For example, in the voter model, an agent i is randomly chosen at each time step, together with one of its neighbour j, and the agent i then abandons its previous belief takes the opinion of its neighbour j.

Other simple models represent agents' beliefs as continuous variables that take the value of a real

number. These are exemplified by the classic Degroot model **(Berger, 1981)** and the bounded confidence model **(Rainer & Krause, 2002)**. The models of this type usually updates agents' beliefs using a weighted average of each individual's belief and those of their neighbour(s). For example, in the bounded confidence model, an agent i is randomly chosen at each time step t, together with one of its neighbour j, who hold the beliefs $\sigma_i(t)$ and $\sigma_j(t)$, respectively. When the condition $|\sigma_i(t) - \sigma_j(t)| < \epsilon$ is met, the agent i updates its belief according to $\sigma_i(t+1) = (1-\alpha)\sigma_i(t) + \alpha\sigma_j(t)$, and the agent j according to $\sigma_j(t+1) = (1-\alpha)\sigma_j(t) + \alpha\sigma_i(t)$. In other words, when the selected agents hold similar enough beliefs according to a threshold ϵ, they update their beliefs as the weighted average of their neighbour's and their own previous beliefs based on a convergence parameter α.

More sophisticated models recognise the distinction between agents' private beliefs and public actions, modelling the former as a continuous variable and the latter a discrete one. The Continuous Opinions and Discrete Actions (CODA) model **(Martins, 2013)** and the more recent social network opinions and actions evolutions (SNOAEs) model **(Zhan et al., 2022)** are examples of this category. To take the CODA model as an example, it models private beliefs as $P(A) \in [0, 1]$, the probability of the A being the best alternative, which is also the probability of an agent publicly displaying the action A. Agents have no access to other agents' private beliefs but only their public actions, which is used to update one's own private beliefs via the Bayes' theorem. For example, upon observing a neighbour j displaying the action A at the time step t, an agent update its private beliefs according to the rule $P_{t+1}(A) = P(A|a_j = +1) \propto P_t(A)P(a_j = +1|A)$, where $P(A|a_j = +1)$ denotes the probability of A being the best action conditioned on the neighbour j displaying A, and $P(a_j = +1|A)$ the probability of the neighbour j displaying A conditioned on A being the best action (which is modelled as a constant).

It is important to notice that there is a recent trend to build psychologically realistic models to represent belief and behaviour dynamics among a group of agents **(Duggins, 2017; Gavrilets, 2021; Tverskoi et al., 2023)**. Still assuming each agent's beliefs and actions are affected by their neighbours, these model further consider the influence of social norm perception, the psychological tendency to conform to peers, external authorities, as well as material cost-benefit considerations **(Gavrilets, 2021)**. To incorporate these psychological and social complexities, a recent unifying modelling framework **(Gavrilets, 2021)** not only considers the distinction between agents' private beliefs and public actions, but also model agents' perception of others' private beliefs and public actions as two separate variables. This allows the representation of various psychological factors in updating beliefs and actions, including cognitive dissonance in taking a private belief (i.e., the aversion towards misaligned belief and action), social projection in perceiving others' private beliefs (i.e., the tendency to project one's own private belief onto others), and compliance with authority.

2.1.2 Model Applications

Models of belief and action dynamics have been applied to studying multiple themes associated with opinion change in a population. These include but are not limited to consensus reaching and polarisation **(Acemoglu & Ozdaglar, 2010; Hassani et al., 2022; Jager & Amblard, 2004; Duggins, 2017)**, the spread of (mis)information **(Watts, 2002; Watts & Dodds, 2007; Pilditch et al., 2022)**, and echo chamber formation **(Madsen et al., 2018; Fränken & Pilditch, 2021)**. Applying mathematical or computational models, researchers are mostly interested in the assumptions and conditions under which certain phenomena can arise.

For instance, a study reviews a series of mathematical models to answer the question about consensus reaching, information aggregation, and the spread of misinformation. ***More specifically, it answers the question about under what conditions 1) agents can hold beliefs in an agreement even when they start with different views, 2) information can be aggregated in the population so that agents holding incorrect beliefs can end up with correct beliefs, and 3) prominent agents can spread misinformation. The study considers both Bayesian (i.e., agents update their beliefs via the Bayes' theorem) and non-Bayesian models (i.e., agents update their beliefs via other rules) (consider remove)*** and concludes that in both types of models there is a tendency towards reaching consensus, but agents' beliefs are not effectively corrected through information aggregation. Misinformation spreads with limited extent in both types of models, which is due to the limited influence of misinformation on Bayesian agents in Bayesian models, and the lack of persistent disagreements in the population in non-Bayesian models.

In addition to belief updating mechanisms, other research has also studied how network structures can influence the emergence of phenomena related to opinion change. ***Focusing on information cascade, a phenomenon where a group of individuals make the same decision sequentially,*** a simulation study finds that the distribution of the size of cascades depends on the connectivity of the network, so does the type of agents who can easily trigger a cascade **(Watts, 2002)**. A subsequent simulation study further points out that influential individuals, i.e., those who have the highest number of connections to others, are not sufficient for triggering cascades, but the existence of a group of other individuals who are easily influenced by the influential ones is critical **(Watts & Dodds, 2007)**.

2.2 Agent-Based Models of Pluralistic Ignorance

There has been a growing interest in using ABMs to study the conditions under which the phenomenon called pluralistic ignorance (PI) can arise. This line of research usually defines PI as the situation where the majority of people in a population express opinions different from their private beliefs **(e.g., Seeme et al., 2016; Wang et al., 2014; Ye et al., 2019)**. The earliest work primarily focuses on the effects of the properties of the network on holding inconsistent private beliefs and expressed opinions **(Centola et al., 2005)**. The research inserts, as an initial condition, a

few agents who hold a private belief held by few individuals in the population (referred to as "true believers") and enforce other agents to adopt the same belief. It reveals that holding inconsistent private beliefs and expressed opinions (which is equivalent to enforcing the belief on other agents in the context of this research) cannot spread widely in the population if the population is fully connected, if the true believers are scattered in the population rather than clustered, or if ties in the network are randomly rewired, breaking the ties that originally exist between local neighbours.

Recent studies start to include in their models psychological processes that have been theorised by research in social psychology. Two processes that have been mostly attended to are social conformity and cognitive dissonance **(e.g., Wang et al., 2014; Seeme, 2019)**. For instance, a research models the change in agents' expressed opinions as a result of the influence from the opinions in the groups formed among the neighbours, representing people's psychological tendency to conform to the group **(Wang et al., 2014)**. Following the cognitive dissonance theory, the model asks the agents change their private attitudes only when the group influence is moderate. When the group influence is strong, the model represents the situation where people are aware of the groups influence and doesn't ask agents to change their private beliefs. Under these assumptions, the research finds that there is a widespread inconsistency between agents' private beliefs and expressed opinions, that is, pluralistic ignorance exists by definition.

Another example is a study that considers social conformity and cognitive dissonance assuming the rationality of agents **(Seeme, 2019)**. It models the agents as having a utility function that is proportional to the rewards from high conformity with its opinion group (expressing similar opinions with the mean opinion of the group) and the rewards from low cognitive dissonance (holding similar private beliefs and expressed opinions). Agents update their private beliefs and expressed opinions by maximising the value of the utility function. The study finds that different degrees of pluralistic ignorance arise under different conditions. When all agents make the update simultaneously, the more opinion groups in the population, the more agents hold inconsistent private beliefs and expressed opinions; when the updating happens sequentially (i.e., agents update one by one), all agents end up holding inconsistent private beliefs and expressed opinions.

2.3 Existing intervention strategy

Discuss existing social norm intervention that corrects misperception

2.4 Current Research

what's new here

ABMs are a type of computational model that simulates in a synthetic environment the actions and interactions of autonomous agents, whose behaviours determine the evolution of the entire system **(Bandini et al., 2009)**. ABMs have been deployed in research in a wide range of fields including climate science **(Simmonds et al., 2019)**, ecology **(McLane et al., 2011)**, epidemiology **(de Mooij**

et al., 2022), economics **(Grazzini & Richiardi, 2015)**, finance **(Bonabeau, 2022; Chen et al., 2017,)**, and social sciences **(Epstein & Axtell, 1996; Schelling, 2006)**. Recently, findings from psychological and cognitive science have also been integrated in building more psychologically realistic agent-based models to study political opinion dynamics **(Duaggins, 2017)**, the formation of echo chamber **(Fränken & Pilditch, 2021; Madsen et al., 2018)**, and interventions to the spread of misinformation **(Pilditch et al., 2022)**. These recent developments exemplify the ability of ABMs in modelling a system where multiple individual-level processes, such as interpersonal influence, social conformity, and commitment to previous beliefs, together influence the emergent properties of the system. Since the system of WWOH actions is presumably such a system, ABM is able to model it (???).

3 Methodology (2000-3000)

4 Results (2000-2500)

5 Discussion and Conclusion (1000-1500)

6 Reference

7 Appendices