Explaining vertical gender segregation: A research agenda

Iñaki R. Longarela
Stockholm University, Sweden
UIT The Artic University of Tromsø, Norway

Abstract

This agenda outlines possible routes to pursue an explanation of vertical gender segregation. The analysis emphasizes the expanding opportunities brought about by a combination of Big Data and public policies, like gender quotas, and uncovers important challenges for which possible solutions are offered. Experimental work is likely to remain very useful in the pursuit of answers to this asymmetric gender presence.

Keywords

Big Data, discrimination, free choice, gender roles, gender quotas, networks, research agenda, vertical gender segregation.

Iñaki R. Longarela, Stockholm Business School, Stockholm University, Kräfiket 15, Stockholm, Sweden. Email: inaki.rodriguez@sbs.su.se
Introduction

Gender segregation has been largely documented and, to varying degrees but with no known exceptions, seems to be present in all countries (see e.g. Charles and Grusky, 2004). Its vertical dimension offers a very consistent pattern, characterized by a much larger share of political and economic power under male control. This particular form of gender segregation has, in recent years, attracted a great deal of attention in the design of public policies aimed at reducing such segregation in workforces. As policymakers increasingly attempt to improve workforce equality through legislation, the scientific community should increase its efforts to analyse and inform such policies and their impact.

A research agenda with this goal has a rich toolbox including both vast amounts of data and the necessary machinery infrastructure for its analysis, which together go under the general name of Big Data. Contemporary social scientists thus have at their disposal massive amounts of information on individuals and their interactions. Register data is becoming increasingly available in many countries following the recommendations of the United Nations Statistics Division. Countries like Norway, Sweden and Denmark are well ahead in these endeavours but most countries are following suit, and their statistical bureaus are implementing a continuous recording of comprehensive demographic and socio-economic data on their citizens and, in many instances, their organizations. Furthermore, the arrival of online social networks like Facebook, Twitter and LinkedIn contain information about individuals and their interactions all over
the world, opening enormous possibilities (and ethical concerns) whose boundaries are still difficult to gauge.

At the same time, new hardware and software are being developed at an exponential rate delivering both processing power that meets the needs associated with larger and larger data sets and programming tools and algorithms that enable the examination of demographic information, language use and social interactions.¹

While advances in Big Data empower increasingly thorough analysis of vertical gender segregation, the very same public policies that the proposed research agenda should inform are also enlarging the research toolbox. Indeed, the proliferation of gender quotas in boardrooms, electoral lists and even government cabinets results in an unprecedented set of natural experiments, which are to a large extent underutilized for analysis. Following Norway’s lead, there are now gender quotas for corporate boards in several European countries. In addition, a much longer list of nations in all continents, including quite unusual suspects like Afghanistan, have implemented some version of electoral and parliament quotas.²

The goal of this note is to elaborate a set of directions for future research, and at the same time evaluate the scope of the opportunities outlined above. In this effort, it has been found useful to organize the exposition around the main explanations for vertical gender segregation that have been advanced by previous qualitative work. Such explanations have mainly gravitated around four major hypotheses: Perpetuation of gender roles (see e.g. Acker, 1990; Alvesson and Billing, 2009), discrimination (see e.g. Wright and Baxter, 2000), absence of networks (see e.g. McPherson et al, 2001; Linnehan and Scullion, 2008) and free
choice (see e.g. Hakim, 1991). The arguments will also lean on important contributions to the literature that have taken place over the last decade, including results from field and laboratory experiments.

**Hypothesis 1: Perpetuation of traditional gender roles**

Proponents of this claim argue that gendered expectations and perceptions persist merely as a result of tradition, and while they may not constitute outright discrimination, they do limit women’s choice consciously or unconsciously. For example, Alvesson and Billing (2009) maintain that “the general cultural conceptions and expectations of the sexes are influenced by those left over from a time when the ‘roles’ of men and women were more fixed”. Acker (1990) points out that “gender segregation is an amazingly persistent pattern and that the gender identity of jobs and occupations is repeatedly reproduced, often in new forms”.

Andersen’s (1988) popular distinction between sex and gender has been the major theoretical underpinning of this position: Gender is the “socially learned behaviours that are attributed to masculinity and femininity”, whereas sex must be defined on the basis of purely biological considerations.

An initial problem with this explanation lies in its potential unfalsifiability. The claim takes as given that the forces that have kept women away from leading positions in most societies are explained by nurture rather than nature. Consequently, all content of this hypothesis is exhausted by that initial stand. Perpetuation is something that must logically follow in view of the low presence of women in top positions in most societies over almost all periods of history.
Once it is taken for granted that socially learned behaviors alone drive this disparity at the top, no empirical result can give us a way to reject such claim. However, one could take a step back and interpret the whole premise as a test of the above initial assumption. Although this line of attack pushes the issue into the quicksand of the nature-nurture debate, it has delivered certain success in explaining the gender gap in labour participation. For example, Fernández et al (2004) use World War II as an exogenous shock to such participation in the form of relative variation in male draft in the United States. They find evidence consistent with such an event spurring an intergenerational propagation mechanism that has led to the higher female employment rates observed today.

Recent literature has also produced tools that could be of future use. Fogly and Veldkamp (2011) calibrate with empirical data a theoretical model, containing local information transmission, in a successful effort to explain the historical dynamics of female labour participation. Their model can be applied to a variety of social changes, including vertical gender segregation, in order to possibly evaluate the role of nurture.

Both studies above rely on Big Data in the form of large surveys combining demographic and economic data with political and social attitudes in the US over a long period of time. The current availability of similar data (possibly of higher quality) across several countries in conjunction with, exogenous shocks like the introduction of gender quotas may be exploited in ingenious ways, to apply their analysis to the subject of this research agenda.

Unfortunately, in order to fully exploit these methods it appears to be necessary to have observable high levels of female involvement in positions of power for an
extended period of time and thus, any research agenda along these lines will not bear fruit in the immediate future.

Alternatively, experimental evidence can at least build in a rigorous way the plausibility of this more general version of this explanation. For example, Gneezy et al (2009) implement an experiment in two different societies, one patriarchal and one matriarchal, to conclude that preferences for competition seem to be mainly determined by environmental factors. New experiments may be designed to evaluate the role of nurture with respect to other individual traits that are considered necessary to the profile of those bearing positions of power.

**Hypothesis 2: Discrimination**

This premise maintains that the unequal treatment of men and women in hiring and promotion explains the past and present levels of vertical gender segregation observed across all countries without exception. This is undoubtedly a plausible explanation as shown by a large series of field and laboratory experiments. For example, Goldin and Rouse (2000) examine auditions for a symphony orchestra position and establish that a screen hiding the sex of the candidate raises the probability that a woman is hired. Coffee and Correll et al. (2007) use a laboratory experiment and an audit study to detect the presence of evaluations that may lead to discrimination against mothers in decisions of hiring and promotion.

This latter study is especially interesting since it also gives plausibility to the potential of the discrimination hypothesis to explain the counterintuitive high levels of vertical segregation that have been documented in countries with strong gender-based family policies like Denmark, Sweden and Norway. As
Mandel and Semyonov (2005) contend, legislation that entitles women to long leaves or temporarily short working days may encourage employers to discard female candidates in positions that require costly training. These policies might exacerbate pure discrimination by magnifying the motherhood penalty documented in the aforementioned study by Correll et al. (2007). Evidence consistent with this possibility is presented by Evertsson et al. (2015) in an analysis that employs three large sets of survey data and finds that family leaves in US, Germany and Sweden have a much more negative impact on female career prospects than other types of leaves. In addition, gender-friendly legislation may also raise the incentives for the practice of statistical discrimination since its asymmetric use may increase the perceived expected labour costs of fertile women relative to otherwise identical men.

In any case, the above experimental work has not been used to its full potential, although, as Wright et al (1995) argue, “ideal data for analysing gender discrimination in access to authority would include direct observations of the discriminatory acts.” In particular, audit studies could play an important role in quantifying the scope of statistical discrimination in positions of power, a task, which has not really been pursued yet. Certainly, these methods in isolation may have clear limitations in establishing links with the low presence of women in the high spheres of power and its variation across different countries. However, as these studies proliferate in relevant environments, a future meta-study may help in fulfilling such a goal.

All the above does not imply that Big Data can be disregarded in the context of the discrimination hypothesis. A natural option can extend the approach presented in Wright et al (1995). These authors use self-reported data obtained
through personal and telephone interviews conducted in the early 80’s in seven industrialized countries. They regress a measure of authority at the workplace on a set of control variables, including individual characteristics, and they establish that a gender dummy is always significant. An ambitious analysis along these lines, which would replace a small sample of subjective data with Big Data, is clearly realistic. However, endeavors of this sort will not be free of methodological caveats. Indeed, attributing the potential significance of a dummy variable to a discriminatory environment requires an initial strong assumption, namely that no gender difference in leading positions may arise from any other unobserved/unavailable variable. In particular, equality of interests and preferences of both gender groups would need to be assumed. Thus, this approach has potentially the taste of a tautological scheme, since the very effect that it is being detected may just be a restatement of the phenomena whose explanation is being pursued. It has been argued (see e.g. Bertrand and Hallock, 2001) that these problems may be less severe in the context of top management, since differences in unobservable characteristics for workers in these occupations are likely to be small. However, Adams and Funk (2012), who resort to a large survey of directors, find that women and men in these positions “differ systematically in their core values and risk attitudes”. This lack of observability may be partially addressed by using data of a qualitative nature. Interviews for at least a matched subsample of individuals with the aim of establishing unobserved psychological traits, interests and preferences may attenuate this problem. Alternatively, such traits, including the well-known Big Five, can be inferred from entries in social networks like Facebook, Twitter and LinkedIn or by textual analysis of press releases (see e.g. Kosinkski et al, 2013).
Finally, the natural experiments facilitated by the imposition of gender quotas across a long list of countries should be an important source of potential answers. Casas-Arce and Saiz (2015) rely on the imposition of such quotas on Spanish electoral lists and find that those parties forced to make bigger relative adjustments to meet the obligatory female representation display relative improvements in their results. This pre-quota non-optimality is hardly explained by arguments that disregard the existence of discrimination. Extending this analysis to countries with electoral quotas is a logical step forward. It is tempting to generalize this approach to boardroom quotas by examining profits and returns before and after the corresponding legal change. However, any implications regarding discrimination from this analysis are probably less valid given the long-run nature of many of the firm's decisions, something that makes optimality assessments much more difficult.

**Hypothesis 3: Lack of Networks**

This hypothesis suggests that women’s low presence in positions of professional authority may be explained by their presumed lack of connections and ties, which in theory are essential in order to gain access to professional upper echelons. According to this hypothesis, the supposedly dominant male typically exerts “homosocial reproduction” or “gender homophily” and only trusts those with the same characteristics (see e.g. Morrison and Von Glinow, 1990 and Ibarra, 1993).

It is worth noting that a good deal of the research along these lines has remained at a conceptual level and this is probably responsible for the existing degree of confusion about the precise meaning of this hypothesis. One may understand the
whole premise as suggesting that since it takes a woman to appoint another woman, the initial low female presence in corporate elites results in the observed male dominance in top management. Gender homophily is thus a cause of this female underrepresentation via a direct mechanism, which has the flavour of discrimination (see e.g. McPherson et al, 2001). An alternative interpretation of the network hypothesis may postulate that gender homophily is indirectly responsible since it takes trust and common perceptions to develop the necessary networks for career advancement (see e.g. Linnehan and Scullion, 2008).

Gender quotas probably offer the greatest potential within the context of the current hypothesis, since they offer two scenarios which favour in a different degree the access of women to social and informal networks and the creation of a female critical mass. A first test may rely on the direct mechanism above and simply estimate relative improvements in the probability of observing women among top managers or politically-appointed senior officials. In doing so, the researcher endowed with Big Data could follow the steps of Bertrand et al. (2014) and extend their analysis to electoral and corporate quotas in different countries. In particular, these authors compare labour outcomes for pre- and after-reform cohorts of Norwegian women, whose qualifications matched those of contemporary members of boards, but who do not occupy board seats. Of course, results that might indicate, as in the aforementioned study, that women are no more likely to be found in positions of power after the mandate, do not represent the end of the story. On the one hand, concerns about tokenism, even in the context of these drastic reforms, may survive, since as Bertrand et al (2014) argue firms “may have decided to ‘game’ the reform by strategically
appointing sub-par women to their boards”. It is thus important to examine, as these authors do, if the newly arrived women have average observable qualifications that match or even improve upon those of women appointed before the reform. On the other hand, it is necessary to examine the indirect network channel since quotas may have failed to improve the network capital of women during the *post-event* window.

In order to proceed with the latter, a second step may construct individual measures of connectedness by using interlocking board directorships, as in Renneboog and Zhao (2011), or crossing paths in employment, education and nonprofessional activities, as in Lalanne and Seabright (2011). This would not only allow for an examination of variations in female network capital after the imposition of the quotas but it would also help in breaking down any detected improvements in female presence at the top between a direct and an indirect channel. Nevertheless, this approach suffers from two major caveats. First, it only identifies networking opportunities as opposed to actual use of such connections. Second, it fails to capture the *strength* of those ties. As Seabright (2012) argue, following the logic of Granovetter (1973), it makes sense to expect that access to top positions may benefit to a much larger extent from acquaintances (weak ties) than from relatives and friends (strong ties). Although the mining of online social networks may come at the rescue, as a growing literature presenting important advances certifies, the above two problems represent important challenges for future research.
Hypothesis 4: Women’s free choice

The assertion that men and women have different attitudes and/or interests that drive vertical (or horizontal, for that matter) job segregation constitutes the fundamental underpinning of this hypothesis. According to this view, the gender gap may be due to differences in traits of a biological nature or differences in preferences, which in turn may be either socially learned or inherited. Field and laboratory work have played an important role in supporting the plausibility of this case. For example, Benbow et al. (2000) find clearly distinguishable patterns of life priorities between men and women, wherein the former seem more likely to pursue status at the expense of family. More recently, Flory et al. (2015) conduct a job-entry experiment in six major US cities and find evidence indicating that women “disproportionately shy away from competitive work settings.” This study is particularly interesting because, as opposed to most of the field evidence in the form of audit studies that analyse the demand side of labour markets, it focuses on their supply side. Further research may continue this direction. For example, direct observation of behaviour that would be very useful in connection with the free-choice hypothesis could also rely on data from headhunting agencies. Their data could be used to establish the probability of individual’s gender affecting their acceptance or refusal of job offers for positions of authority.

Note that this hypothesis has also explanatory potential in unsuspected territory since it is again possible to turn the puzzle on its head and speculate that women-friendly policies aggravate rather than attenuate women’s underrepresentation in top management. If women are less careerist than men,
and more prone to pursue a good work-life balance as evidence from Benbow et al. (2000) seems to suggest, then by lowering the short-term costs of long parental leaves and part-time work, these policies may increase the number and length of women’s career interruptions, which may in turn impede their advancement up the corporate ladder in the long run. An identical outcome may also be the result of a decision driven by monetary incentives rather than interests or career attitudes. The reasoning in this case is of a Beckerian nature. An imbalanced breadwinning scheme, where men are responsible for a larger chunk of total income, may also lead to a gender subdivision of tasks whose potential harm to women’s careers may be exacerbated by welfare policies of the exact opposite intent.

There is an alternative channel by which women’s free choice may drive this low female presence. The core of this argument, which to the author’s knowledge is novel, appeals to the relatively lower economic incentives for a corporate career associated with a more egalitarian wage structure. Indeed, in 2007 top executives in the US earned about 344 times more than the average salary (Landy, 2008), whereas the corresponding figure for a country like Norway was approximately 8 (Randøy and Skalpe, 2009; Statistics Norway). Furthermore, countries like Sweden and Japan, where CEO compensation is low relative to other OECD countries (Abowd and Kaplan, 1999), display very high levels of vertical gender segregation (Wright et al., 1995). Hence, those individuals who are already less enthusiastic about highly competitive environments and inflexible working hours may find the executive office even more unattractive. This speculation could be pushed further from a slightly different angle. The same lower incentives may follow if all individuals are entitled to health and
social welfare amenities, and if most individuals (even outside of executive positions) earn a salary sufficient to consume high-status goods. Additionally and by the same token, evenly spread economic growth may aggravate the gender gap in top management presence.

Whichever the channel chosen, a serious challenge emerges since the problem of the unobservability of preferences is aggravated by the fact that female absence at the top cannot be used as an instance of revealed preference, as after all, that is exactly what one is trying to test.

Certainly, at first sight one can be seduced by an apparently convincing argumentation. If any logistic regression using individual register data gives a significant value to the wage variable suggesting that the probability of finding a female executive increases with its retribution, this can be taken as evidence supporting this hypothesis. Clearly, it is improbable that the intensity of discrimination is positively correlated with the wage variable or, in other words, the thickness of the glass ceiling is unlikely to shrink when the reward increases. On the other hand, if the wage variable is not significant, this could constitute evidence against the hypothesis. It is true that there are other variables that may shape, to a large extent, women’s preferences and decisions and, in particular, civil status and motherhood are reasonable candidates. However, if the true barrier to the executive room is a free decision (driven by interest, not talent), it is hard to argue that monetary incentives are irrelevant to such a decision once the above individual variables are appropriately controlled for.

There are two serious problems with this line of reasoning. First, the possible presence of other relevant unmeasured differences needs to be acknowledged, which although dismissed by Bertrand and Hallock (2001), have prevalence that
is supported by recent evidence (Adams and Funk, 2012). Big Data may partially address this first issue. An even more serious concern may be a selection-bias problem caused by the potential existence of gender discrimination. If this barrier is present, those women who in the end make it to the top are exceptional to a degree that may justify their potentially higher wages.

**Discussion**

Vertical gender segregation is a puzzle which has a great deal of subtlety and whose full explanation may prove to be a difficult goal. Big Data and institutional changes like the implementation of gender quotas could not only help reveal the relative importance of the previously postulated explanations, but also uncover the hidden mechanisms that govern this empirical regularity. Within the latter, there is an important issue that remains relatively elusive and which might be important to settle in the near future so that research can more effectively focus on the former. Just like the explosion of labour participation rates among childbearing mothers has been the major driving force of the closing of the gender gap in employment in developed countries, there is important evidence indicating that motherhood is the major explanatory variable of the wage gap (see e.g. Bertrand et al., 2010; Correll et al. 2007). Determining empirically how motherhood affects the probability of finding a woman in positions of power demands attention and it is increasingly feasible with an intentional research agenda exercised with the rich toolbox that social scientist have at hand nowadays.
It is also important that efforts to address the puzzle of women’s underrepresentation at the top are efficiently invested in, and avoid the (highly ideological) mud of the nurture-nature debate will likely accelerate the search for reliable answers. For this purpose, the classical divide between glass-ceiling explanations versus women’s free choice arguments may have to be replaced by a more constructive categorization. In particular, a general statement of the free choice hypothesis, as the one chosen in this note, is compatible with highly deterministic views about the role of nurture in shaping gender roles and individual preferences and decisions. These two hypotheses are in any case unlikely to be empirically distinguishable when it comes to examining actual labour market outcomes. It might thus be found desirable to treat them as equivalent explanations.

Nevertheless, even though Big Data will present clear opportunities in this research quest, experimental evidence is likely to retain a very important role, especially, in handling the observability problems that similarly affect the testing of the discrimination and free-choice hypotheses.

**Acknowledgments**

I would like to thank three anonymous referees and Prof. Paul Brook in his editorial capacity for their suggestions and very valuable insights. I am also grateful to my colleague, Prof. Lene Foss of UIT The Arctic University of Norway, for her contributions to previous versions of this paper. My most especial thanks go to Prof. Claire Armstrong, Dr. Carolina Silva, and Mallory Imler Powell, who read several versions of this paper and provided very useful comments.
Notes

i New full-blown computer languages like Python offer features favoring the development of Application Programming Interfaces (API’s) and the creation of fundamental tools for textual analysis, web scrapping and data mining of social networks (see e.g. Russel, 2011). Also, the sophistication of algorithms with this aim is increasing at an astonishing rate (see e.g. Gupte and Eliassi-Rad, 2012).

ii See the information compiled by the Quota Project at http://www.quotaproject.org for details.

iii See e.g. the UN’s International Labour Office statistics at http://laborsta.ilo.org.

iv See e.g. Sohrabi and Akbari (2016) and the references therein.

References


Russell MA (2013) *Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More*. O'Reilly Media.


Iñaki R. Longarela is an Assistant Professor at the Stockholm Business School within Stockholm University and a Visiting Associate Professor at UiT-The Artic University of Norway. Within sociology, his research interest focuses on gender segregation issues.