Documento de Trabajo - 2018/06

Women across Subfields in Economics: Relative Performance and Beliefs

P. Beneito
(University of Valencia and ERI-CES)

J. E. Boscá
(University of Valencia and FEDEA)

J. Ferri
(University of Valencia and FEDEA)

M. García
(FEDEA)

fedea

Women across Subfields in Economics: Relative Performance and Beliefs

P. Beneito
J. E. Boscá
J. Ferri
M. García

¿Por qué la Economía está consolidándose como una disciplina con escasez de mujeres? Esta pregunta, aun no siendo nueva, está siendo objeto en los dos últimos años de una gran atención por parte de la profesión, especialmente en la academia. El bajo porcentaje de mujeres en la investigación académica en economía es un hecho desde hace décadas, pero, al contrario de lo que podría esperarse, dicha escasez relativa no parece haber evolucionado positivamente en los últimos años. Las tasas se sitúan entre el 20 y el 25 por ciento, dependiendo de la base consultada.

Relacionado con lo anterior, la tasa de estudiantes mujeres de economía se sitúa entre el 25 y el 40 por ciento, dependiendo del país considerado, y con tendencia a la baja. Los motivos por los que se argumenta que esto puede estar ocurriendo son variados y recurrentes, si bien la evidencia existente al respecto es todavía limitada. Algunas de las explicaciones más comunes son que: i) las mujeres no optan por carreras de contenido matemático (algo desmentido actualmente por las tasas crecientes de mujeres en matemáticas); ii) las mujeres no encuentran la economía interesante (a juzgar por encuestas pasadas a estudiantes pre-universitarios que han seguido cursos de economía en bachiller); iii) el estereotipo del economista como una figura de varón.

Las consecuencias de esta escasez relativa de mujeres en la profesión no se reducen, aun siendo muy importante, a una cuestión de igualdad de género. Una primera línea de investigación reciente pone de manifiesto que los puntos de vista sobre temas centrales en economía pueden mostrar diferencias considerables por género. Esto, junto con el hecho de que las mujeres investigadoras en economía tienden a concentrarse en ciertos temas de investigación más que en otros, puede acabar generando un sesgo bastante marcado en los temas que reciben mayor atención, y en el tipo de respuestas que reciben. Una segunda línea actual de investigación ha puesto el énfasis en la idea de que la investigación se enriquece con la diversidad dentro de los grupos que la llevan a cabo, siendo el género una de las dimensiones más destacadas de dicha diversidad.

Nuestro trabajo aporta una explicación nueva al escaso atractivo relativo que la economía parece ejercer sobre nuestras jóvenes estudiantes. En resumen, obtenemos que las áreas de investigación económica donde las mujeres tienen más presencia académica se relacionan con las áreas de estudio donde las mujeres estudiantes en la universidad obtienen mejores calificaciones en relación a los hombres. Sin embargo, justamente estas áreas son las menos conocidas por nuestros estudiantes, que tienden a asociar la mayoría de los trabajos de un economista con un macroeconomista, mayoritariamente varón.

Esta explicación surge a partir del análisis que efectuamos en tres frentes. En primer lugar, utilizamos técnicas de extracción de información web aplicadas a los trabajos presentados en el congreso anual de la American Economic Association durante años recientes. A partir de la clasificación de los trabajos según sus códigos JEL, observamos que el porcentaje de mujeres investigando en economía está en torno

al 24 por ciento, se mantiene bastante estable a lo largo de los años analizados, 2010-2016, y con una distribución desigual por áreas: cerca de un 30 por ciento en microeconomía y en torno al 18 por ciento en macroeconomía, finanzas y econometría/matemáticas. Analizando los abstracts, temas específicos como economía de la educación, laboral, de la salud, o de género presentan cerca de un 45 por ciento de mujeres, frente a un 13-15 por ciento de mujeres en temas específicos de finanzas y macroeconomía.

En segundo lugar, y constituyendo evidencia inédita al respecto, estimamos el diferencial de notas de las mujeres frente a los hombres en sus estudios universitarios de economía por áreas. Utilizando datos de registro de la Universidad de Valencia, analizamos las calificaciones obtenidas por hombres y mujeres matriculados en economía desde 2010 a 2014. Disponemos de sus notas en todas las asignaturas cursadas durante la carrera, que clasificamos en distintos campos de estudio (finanzas, macroeconomía, microeconomía, asignaturas instrumentales, y 'otras'). Controlando por su capacidad académica (notas de entrada), características del individuo como edad, situación laboral y lugar de residencia, así como por nivel educativo y tipo de ocupación de padre y madre, obtenemos claras diferencias entre hombres y mujeres en las asignaturas de macroeconomía y microeconomía. En nuestra estimación en asignaturas optativas, por cuantiles de notas, y después de aislar el efecto que pueda tener la manera en que los estudiantes realizan su selección de asignaturas, obtenemos que las mujeres superan en notas a sus compañeros varones en microeconomía, mientras que sus notas son inferiores en las asignaturas de macroeconomía. Los resultados son más abultados en los cuantiles altos de la distribución de notas.

Por último, a partir de una encuesta pasada a nuestros estudiantes de economía durante el curso 2017-2018 analizamos la información que nos proporcionan acerca de su visión de distintas asignaturas y de la propia profesión. Los resultados más llamativos se refieren a la concepción equivocada que mantienen acerca de la profesión: preguntados acerca de ocupaciones propias de un economista (trabajar en un centro de estudios/investigación, servir de asesor a instituciones públicas/privadas, hacer carrera académica, hacer oposiciones para la administración, trabajar en el sector privado) contestan mayoritariamente que los asocian a la macroeconomía, y también, en algunos de los casos, en mayor medida a hombres. Además, en mayor proporción que los hombres, las mujeres declaran encontrar la microeconomía más intuitiva, más fácil de aprobar, más preocupada por los problemas de la sociedad, más interesante en general, y declaran obtener mejores notas. También observamos que los estudiantes que declaran haber entrado con vocación en la carrera manifiestan una mayor preferencia por la macroeconomía. Por último, del estudio de las interacciones sociales dentro del aula encontramos que las mujeres son especialmente influenciables por sus compañeros varones en las asignaturas relacionadas con la macroeconomía, algo que no sucede con las asignaturas de corte más microeconómico. Esta pauta de influencias sociales es diferente de la que encontramos para los estudiantes varones, a los que sólo parece afectar el comportamiento y las preferencias de sus compañeros varones.

En definitiva, en nuestro estudio evidenciamos fallos de información por parte de los estudiantes sobre los temas a los que se dedica la economía y las salidas profesionales de las distintas áreas que configuran la disciplina. La información que llega a los estudiantes está sesgada justamente en contra de aquéllas áreas en las que comprobamos que las mujeres tienen ventaja comparativa sobre los hombres en términos de sus resultados académicos y del interés que les despiertan ciertos temas. Por el contrario, los estudiantes confieren a la macroeconomía un protagonismo dentro de la

economía que no se corresponde con su verdadero peso en el conjunto de trabajos académicos, y observamos que los estudiantes varones tienden a superar a sus compañeras mujeres tanto en resultados como en la influencia que ejercen en las asignaturas de macroeconomía. No es de extrañar, por lo tanto, que la combinación de creencias, intereses y resultados provoque que las mujeres se sientan poco inclinadas por los estudios de economía.

Women across Subfields in Economics: Relative Performance and Beliefs

P. Beneito[a,b] , J.E. Boscá[a,c], J. Ferri[a,c] and M. García[c].

This version January, 2019

Abstract

The relative scarcity of female students enrolling in economics has become entrenched over the last decade. We provide evidence of gender differences in performance and in preferences across subfields of the discipline and explore students' beliefs about the profession and their opinions on different subjects. The areas where women stand out relative to men are those that seem to be less well known to our students. We work on three fronts. First, using web scraping and machine learning techniques, we document the relative presence of women across subfields in recent AEA annual meetings. Macroeconomics and finance register the greatest scarcity of women. Second, using administrative records for economics students in a large public university in Spain from 2010 to 2014, we find that women outperform men in microeconomics, while men outperform women in macroeconomics, more evidently in the upper tail of the grades distribution. Finally, data gathered through a self-statement survey given to economics majors reveal that (i) they hold a macroeconomics-biased view of the economics profession; (ii) they exhibit gender differences in their perceptions of the interest and difficulty inherent in different subfields (macro vs. microeconomics); and (iii) their interests and performance are influenced differently by their male and female peers in macro and microeconomics subjects. Taken together, these three pieces of evidence provide a plausible explanation as to why women are relatively less attracted than men to economics, and suggest lines of action to redress the imbalance.

Keywords: Gender, Economics subfields, Information and Beliefs.

JEL classification: A22, J16, J82, D91.

Corresponding author: francisco.ferri@uv.es

[a] University of Valencia, Spain; [b] ERI-CES; [c] Fedea.

We acknowledge the University of Valencia (UV) for providing the administrative data, as well as our collegues at the Department of *Análisis Económico* for their collaboration in gathering the survey data. We also thank Antonio Cabrales, Laura Hospido, Nagore Iriberri and Virginia Sánchez Marcos for useful comments. Pilar Beneito acknowledges financial support from the Spanish Ministerio de Economía y Competitividad (ECO2014-55745-R). José E. Boscá and Javier Ferri acknowledge financial support from Ministerio de Economía y Competitividad (ECO2017-84632-R), Prometeo (GVPROMETEO2016-097), Fundación Rafael del Pino and BBVA-Research.

1 Introduction

Why are there so few women in economics? This is a question that has captured the attention of economics scholars over the last few years, and one for which the profession does not have a widely accepted answer. In this paper we look inside the discipline and explore whether women's and men's relative performance and their preferences for different subfields of economics, be they natural or socially-shaped, can help identify the areas that may be more congenial to women. Linked to this, are young students aware of the rich variety of topics covered by different subfields of economics? If the information that reaches them is skewed against fields where female students feel more comfortable, then we cannot expect women to feel particularly attracted to economics.

We provide evidence of gender differences in performance and in preferences across subfields of the discipline and explore students' beliefs about the profession and their opinions on different subjects. The areas where women relatively stand out are those that seem to be less well known to our students. We work on three fronts. First, looking at women in academia, we document their revealed preference for specific research topics as compared to men's preferences. Second, we analyze the degree to which the relative performance of women and men across subfields of economics exhibits differences even at undergraduate level. Finally, we explore students' beliefs about the profession and their opinions on different subfields, and relate them to gender interactions in the classroom.

Three complementary sources of data are used to uncover these patterns. First, using web scraping and machine learning techniques, we follow women's attendance at the annual AEA meetings in recent years. We document, first, the significant scarcity of women in the economics academic profession, and, second, the most remarkable gender differences in specific subfields. In particular, session JEL codes from 2010 to 2016 show that women account for barely 24 percent of all authors, with macroeconomics and finance being the subfields with the most severe underrepresentation, at around 18 percent on average over the period. Additionally, we dig into the abstracts to classify the papers by topics, finding sizeable differences in terms of female representation.

Second, we study if the observed differences between subfields are already being established in the early stages of the study of economics. Using administrative records for students of economics in a large public university in Spain (University of Valencia, UV henceforth), we analyze grades obtained by students from 2010 to 2014 in different courses of the economics major, which we classify into subfields (macroeconomics, microeconomics, finance, methods and other). Our econometric results from quantile regressions indicate that women outperform men in microeconomics-related courses, while men outperform women in macroeconomics-related ones; in addition, these relative differences become more noticeable in the upper tail of the grades distribution. No remarkable differences are found in the

rest of the subfields. These results are shown for both compulsory and optional courses, in the latter case after controlling for students self-selecting into certain courses, by means of semi-parametric selection controls.

Why do these differences appear? Since in our estimation we control for students' overall academic ability as well as their specific methodological ability, we can rule out possible gender-based differences in mathematical and analytical ability as a reason for the observed differences across subfields.

Finally, intrigued by these results revealing gender differences in performance in macroeconomics vs. microeconomics-related courses, we analyze information from a specifically-designed self-statement survey given to students enrolled in the economics major at UV in 2018. Both their direct answers as well as our regression results with these data reveal that: (i) students hold a macroeconomics-biased view of the economics profession, often mistakenly associating many economists' occupations with macroeconomics, and those women who view economics as a macro and male-dominated profession tend to report a worse performance in macroeconomics; (ii) self-defined vocational students (those for whom economics was their first best choice among university degrees) are more likely to report a higher relative preference for macroeconomics; (iii) performance in microeconomics is appreciably affected by how interesting students find this subfield, with more women than men reporting a greater interest in microeconomics than in macroeconomics; (iv) students exhibit gender differences in their preferences for the two subfields and in their self-reported relative performance that are in line with the observed distribution of women across areas of research and with the results obtained with the administrative data; and (v) both students' individual interests and their performance seem to be influenced by their male and female peers' interests and performance, with these influences exhibiting differences in macro and microeconomics courses; among other results we find that, in macroeconomics, women's interests respond to their female peers' interests in this subfield while, in terms of marks, only male students exert spillover effects on their peers' performance (both women and men) in macroeconomics classes.

Taken together, the three pieces of evidence provided in this paper suggest a subject-bias in terms of visibility, performance and perceived interest across subfields, which may be connected with the declining trend in female students choosing to study economics.

The rest of the paper is organized as follows. In Section 2, we frame our paper in the ongoing debate about the scarcity of women in economics and highlight its contribution to the literature. Then, in Sections 3, 4 and 5, we present the results from web scraping, administrative data, and economics students' self-reports, respectively. Finally, in Section 6, we summarize our main findings and conclude.

2 The ongoing debate and our contribution

In the last few years, a growing number of voices in both academia and the mass-media has been reflecting on the situation of women in the economics profession.[1] However, despite the recent resurgence of the issue, the underrepresentation of women in economics is not a new concern. The Committee on the Status of Women in the Economics Profession (CSWEP) of the American Economic Association (established in 1971) and the RES Women's Committee (established in 1996) have been working to promote the role of women in economics for some time now. In fact, authors such as Ferber (1995), more than two decades ago, pointed out that the low representation of women and minorities among students of economics had been noticeable for some time. The author reported that around 19 percent of PhDs in economics were earned by women in 1988-89, but ended with the optimistic prediction that "*(...), there is reason to believe that in economics, as in most other disciplines, women's progress will eventually accelerate*", (p-357).

That does not appear to have been the case. On the contrary, according to statistics on the published economists registered with the RePEc Author Service, female representation in economics academia (as of May 2018) still stands at a worldwide average of around 19 percent.[2] At the undergraduate level, worldwide figures indicate that around 25 to 40 percent of those opting for economics majors are women, with a declining trend over the last 10 years. For the US, as an example, Wolfers (2018)[3], using data from the Census Bureau's American Community Survey, reports that women made up 35 percent of economics majors in 2016, a figure about the same as in the early 1980s. Bayer and Rouse (2016), using data from the National Center of Education Statistics in the US between 1995 and 2014,

[1] In the summer of 2017, Alice Wu, caused a sort of commotion in the profession revealing the sexist comments posted on the online forum Economics Job Market Rumors (Wu, A. H., 2018, Gendered Language on the Economics Job Market Rumors Forum, in *AEA Papers and Proceedings*). The AEA President, Alvin E. Roth, charged an ad-hoc committee to evaluate various aspects of professional conduct and diversity in economics. The list of posts and articles appearing since then is becoming countless. Just to mention some examples: Justin Wolfers (University of Michigan) in *The New York Times* (https://www.nytimes.com/2018/02/02/business/why-womens-voices-are-scarce-in-economics.html, February 2, 2018); Sarah Smith, (University of Bristol, Chair of Royal Economic Society (RES) Women's Committee) in (https://www.bristol.ac.uk/efm/news/2018/still-a-long-way-to-go-a-century-on.html February 6, 2018); Anne Boring (Universiteit Rotterdam) and Soledad Zignano (Banque de France) in the *Banque de France, Eco Notepad* (https://blocnotesdeleco.banque-france.fr/en/blog-entry/economics-where-are-women July 3, 2018); Gemma C. Tetlow (UCL) in *Financial Times*, (https://www.ft.com/content/0e5d27ba-2b61-11e8-9b4b-bc4b9f08f381 April 12, 2018); Soumaya Keynes' TEDx Talk ("Why Are There So Few Women In Economics?": https://www.youtube.com/watch?v=Heiu7RPVggw, July 27, 2018). Other experts' concerns and opinions, for example, on the BBC-Radio, (http://www.bbc.co.uk/programmes/b01875r3/topics/Economics, Claudia Goldin (Harvard University) May 1, 2018, Stephen Machin (CEP and LSE), May 8, 2018, Beatrice Cherrier (THEMA), May 15, 2018.

[2] https://ideas.repec.org/top/female.html, last accessed May 23, 2018.

[3] *The New York Times* (https://www.nytimes.com/2018/02/02/business/why-womens-voices-are-scarce-in-economics.html, February 2, 2018).

calculate the percentage of women with a bachelor's degree in economics in 2014 at 28.4, a lower representation than in other social sciences, business and management, and even in STEM (science, technology, engineering and math) fields. In the UK, the Universities and Colleges Admissions Service (UCAS) cites percentages below 30 percent for female undergraduate students of economics in 2013 (after a decade of decline). In the university that is the focus of this study, UV in Spain, women enrolled in the economics major in 2017 account for 36 percent of the total.

The study of Amanda Bayer and Cecilia Elena Rouse (2016) provides a complete account of the negative consequences that the under-representation of women may have in the economics profession, as well as the possible reasons for the current gender imbalance. As the authors stress, fairness, albeit highly important, is not the only reason to care about the gender balance. They survey the recent strands of research that posit diversity within a profession as a necessary factor to ensure the highest quality of knowledge. One of the strands focuses on the difference of opinions across different groups. In economics, May et al. (2014) find that male and female economists hold different views on important economic outcomes and policies. For example, female economists differ notably from male economists in terms of their preferences about the degree of desirable government regulation, income distribution issues or labor market policies. Hence, the gender imbalance may also imply a sustained bias in the prevailing range of topics and a reduction in the diversity of points of view, which may result in negative effects on research quality.

Another line of research summarized by Bayer and Rouse, mostly based on behavioral analysis through laboratory experiments, shows that diversity changes group dynamics and decision-making. These studies find that, in general, groups with mixed gender composition develop a richer interaction and a "collective intelligence" that helps them to perform better than groups lacking diversity (see, e.g., Woolley et al, 2010, or Hoogendoorn et al., 2013, and references therein).

As regards the possible reasons for the scarcity of women in economics, the authors survey both supply-side and demand-side factors alluded to in the literature. Among the former, we can cite: (i) the recurring explanation that math requirements inhibit women's entry into economics (although there is not in fact strong support for this justification - see Emerson *et al*, 2012 - especially given the fact that there has not been a corresponding decline in the percentage of women enrolling in math degrees); (ii) the role of previous exposure to economics, with more women than men reporting that they do not find economics interesting (Calkins and Welki, 2006, additionally report that perceived interest is a key factor determining degree choices); and (iii) the stereotypical view of economics as a male-dominated field.

On the demand-side, discrimination against women is seen as one of the likely explanations, even if it corresponds to implicit or unconscious bias reflecting existing stereotypes (see

e.g., Greenwald and Krieger, 2006). The fact that the profession is seen as a male-dominated one may have an effect not only on hiring practices, but even possibly on day-to-day interactions such as instructors giving advice to their students. Both female and male instructors may potentially exhibit such biases (Moss-Racusin et al., 2012).[4]

In this paper, we offer a new insight into the possible reasons for the gender imbalance in economics. The suspicion that there may be a subject bias in the profession, along with the likelihood of different gender preferences for different subjects/subfields in economics, provides a plausible explanation as to why women are relatively less attracted than men to economics. Our paper investigates the relative performance of women and men in different subfields of economics, and their different perceptions as regards these subfields and the economics profession. Our main finding is that women display relatively better performance and higher self-selection into microeconomics-related subjects from undergraduate level up, and that students hold a macro-biased (in part also a male-biased) view of the profession. Younger students at pre-university stages cannot be expected to be better informed.[5] Hence, if young female pre-university students are not aware of the diversity of economic topics, including those in which, according to our findings, they show more interest and perform better, then it is hardly surprising that they do not feel inclined to study economics.

There may be several reasons for the lack of information and, in particular, the subject-bias of the information that reaches young students (or perhaps people in general). Topics in macroeconomics subfields are more frequently on the news and also more frequently presented to general audiences, while many other varied and probably more specific topics analyzed by economists are not so directly publicized outside academia. Added to this, during the last decade, the global financial crisis might have helped to accentuate the visibility of some specific topics over others. Over the last 10 years, we have all become accustomed to the presence of a (more often than not male) economist in the mass-media offering an analysis of causes, consequences and predicted evolution of key economic indicators. Current university students, aged from 18 to around 22 years old, were children at the outset of the Great Recession, and have probably grown up envisaging economists in this stereotypical way.

On the other hand, as we show in Section 3 below, there is a great variety of topics that

[4]In a recent study that has received great attention, Sarsons (2017) documents such a bias reflected in the higher penalty for coauthoring suffered by women economists, particularly if they coauthor with men. Hengel (2018) presents evidence that women's writing is held to higher standards in academic peer review publishing. The students themselves may be affected by the male stereotype. MacNell et al. (2015) and Boring et al. (2016) provide evidence that students rate online teachers higher when those teachers use a male name, regardless of their actual gender.

[5]Della Giusta et al. (2017), for a sample of pre-university teenage students in the UK, show how biased their perception is as regards the university majors. Also, an online survey by Economics Network administered to 1,700 UK inhabitants shows that economics is associated with aspects such as predictions, finance and business, to the exclusion of virtually everything else (https://www.economicsnetwork.ac.uk/research/understandingecon, last accessed March 23, 2018.)

can be classified as microeconomics. Apparently we need more 'key words' to illustrate what microeconomics is about than we would to convey a broad but approximate idea of what macroeconomics is about. This poses greater difficulty in convincing potential students as to how appealing (micro)economics can be. Hence, even if "economics is what economists do", it is also the case that "economists do not (only) do what people think they do".[6] If we are not able to communicate the varied and appealing areas economists deal with and, moreover, if this lack of information primarily affects a particular area of economics, then potential students that would otherwise have been attracted to those 'hidden' areas of economics may not enroll in economics majors.

Why we should care about the gender imbalance may in turn depend on the underlying causes of this imbalance. If the fact that women do not feel particularly attracted to economics is caused by an information problem, then such a scarcity may be, in turn, a symptom of other problems in the discipline. Lack of information and/or misinformation for pre-university students choosing among university majors may entail an inefficient allocation of resources, as in many other markets. Talented students may be dissuaded from enrolling in majors such as economics if they misunderstand what economics is really about and the career opportunities that this discipline opens up for them.

An additional consequence of a lack of information on particular areas of analysis within economics is that the general public will also hold not only a subject-biased but also a gendered view of the profession. If the areas of research with a higher proportion of women (see Section 3 below) and where female students perform better (Section 4 below) are less well known to the public (Section 5 below), this may affect further the visibility of female economists. Given the documented effect of female role models on gender imbalances, this may constitute a factor that contributes to sustaining gender inequality.

Moreover, the lower level, and even declining trend, of female enrollment in economics, along with the fact that they seem to self-select more into microeconomics-related areas, could accentuate the gender imbalance between subfields within economics. In some cases, such as macroeconomics and finance, we might end up with an almost entire absence of professional female economists.[7]

In this paper, our aim is not to offer an exhaustive exploration of the causes that may women lead to select specific areas of research or perform relatively better in some subfields, although our survey data offer some clues in this regard. Rather, our primary goal is to document the existing differences and connect them with our students' self-reported information biases.

To our knowledge, our paper is the first to provide the sort of evidence we present be-

[6]The classic quip "Economics is what economists do" is attributed to Jacob Viner.

[7]The striking scarcity of women in some areas complicates the success of some recent initiatives such as Spanish economists and academics boycotting events without female experts. https://sites.google.com/view/nosinmujeres/inicio; last accessed May 31, 2018.

low, particularly in the last two parts. The information collected through web scraping adds to the scarce formal evidence provided by the literature so far as regards the areas of research where women self-select. The most similar work to this part of our paper is the study by Chari and Goldsmith-Pinkham (2017), where the authors provide information on the proportion of women in each subfield at the NBER summer meeting. We extend that information by scraping the website of the wider annual AEA meeting, and by offering detailed information on the percentage of women by research topics. As regards the results with the administrative data, to our knowledge, they represent an entirely new contribution to documenting the relative academic achievement of male and female economics students by subfields. It is particularly striking how our econometric results in this part highlight differences in the academic performance of men and women that line up with the observed worldwide trends in the choice of research areas by gender. Finally, the self-response questionnaire has been designed to find out what might be behind the econometric results with the administrative data. Since it is targeted at students in the same major and university, it uniquely complements those results.

The extent to which our results can be extrapolated to a broader scale is something we cannot categorically state. However, given that gender preferences are probably quite universal, that the curriculum of the economics major at UV is quite similar to that of other faculties and countries, and the fact that we end up observing patterns that are consistent with observations already established in the literature, we believe that our work opens up a new avenue of research that is worth exploring in other contexts.

Our results suggest reasons why women do not feel particularly attracted to economics, and point to lines of action to redress the imbalance. The main takeaway message is that economists have a communication challenge, as many voices have already expressed, and that this challenge may be particularly urgent in the case of pre-university students. Our results further support the view that this problem of information may explain part of the observed scarcity of women in economics; in addition, they point to the need to better publicize specific areas that, although they may well be highly appealing, are as yet the less well known. Our paper, then, underscores the relevance of adopting and extending projects such as CORE[8], to help fill this information gap. The misinformation also affects our students' perceptions as regards the main occupations for economics graduates and the areas of specialization that may be open to them. Economics teachers at both pre-university and university levels need to be aware of the lack of information that affects our students, and should make a conscious effort to alleviate the problem.

[8]The Curriculum in Open-access Resources in Economics (CORE) project, supported by the Royal Economic Society, is an international initiative created in 2013 as a response to concerns that the economics curriculum was becoming disconnected from the needs of learners, economics departments and employers, and also from the lived experience of potential students. https://www.ineteconomics.org/education/curricula-modules/core-project

3 Women in the AEA annual meeting and the variety of topics across subfields

To throw light on the representation of women in economics across different subfields, we look at the scientific programs of the annual meetings of the American Economic Association (AEA) in recent years. First, we classify the papers presented at AEA meetings into broad areas, according to the JEL codes of their corresponding sessions[9]. Second, we dig into the paper abstracts and obtain word clouds using a machine learning approach to classify papers by topics.

3.1 JEL approach

We extract information on the programs of the annual AEA meetings from 2010 to 2016, obtaining the JEL codes of the sessions and the names of the authors. Based on the JEL codes, papers are classified into one of the following five subfields or categories: Microeconomics (including both theoretical and empirical micro-oriented topics); Macroeconomics (also covering international economics); Finance; Methods (encompassing mathematical and quantitative methods) and Other (including a set of miscellaneous categories difficult to categorize elsewhere). To calculate the share of women, we account for the total number of authors of all papers submitted in a particular category and classify them by gender according to their first name. For this purpose, we rely on three different databases. First, we use the first-names database published by the U.S. Social Security Administration, created using data from Social Security card applications. We merge this database with the list of authors from the AEA meetings. Then, we assign an author as female if the probability that the name is female is higher than 0.95, and assume an author is male if the probability that the name is female is lower than 0.05. For the non-classified names, we use the database constructed by Tang et al. (2011), who use Facebook to collect data on first names and self-reported gender, and as before we use probabilities to assign gender as female or male[10]. Lastly, for the names that remain non-classified after these two iterations, we use the names database developed by Bagues and Campa (2018), who assign gender to the different names with a probability of one. Any authors who either fall within the [0.05 0.95] interval or cannot be found in any of the databases are excluded.[11] Our observational unit is typically an author-paper-field-year match (author to simplify). We register 22,609 authors over the

[9]The complete list of these JEL codes and their correspondence with the different areas is available in Table 6 in Appendix A.

[10]This database is also used by Chari and Goldsmith-Pinkham (2017).

[11]After using the three database, the average of non-classified names across all years and areas is 13.8 percent. We assume that the gender distribution by economic subfield of excluded authors is not significantly different from that of gender-identified authors. Results that include a non-classified gender category are qualitatively similar to the ones offered in the paper.

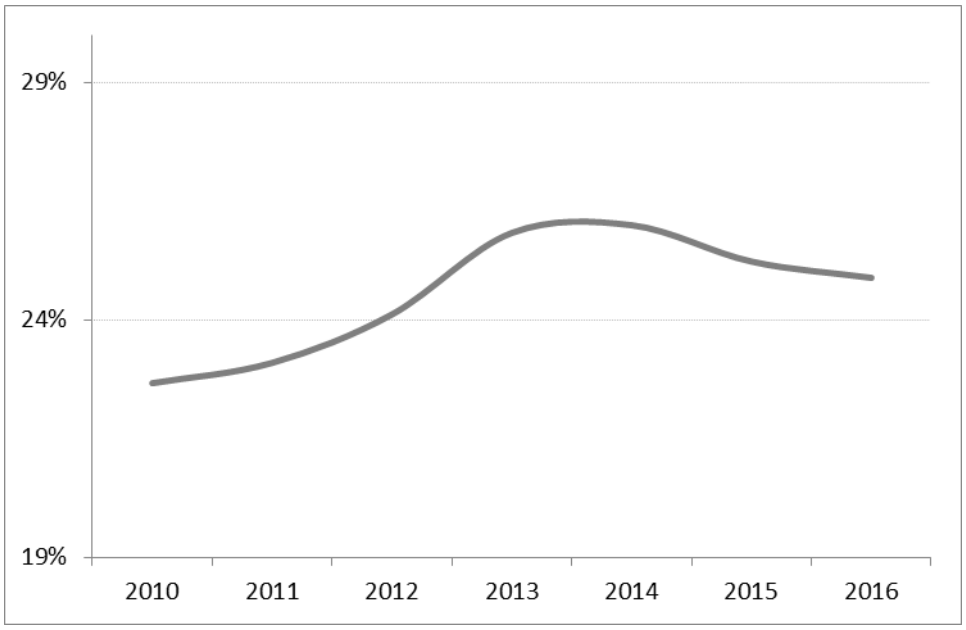


Figure 1 : *Share of women: all research areas (2010-2016)*

whole period, of which 5,611 (24.82 percent) are female. Of course, the same author may appear more than once in the roster.

Figure 1 shows a fairly stagnant evolution of the share of women around the average, with a minimum of 22.7 percent in 2010 and a maximum of 26 percent in 2014. Additionally, Figure 2 represents the evolution of the share of women across areas, where we uncover a clear gap between two research clubs: one related with Macro, Finance and Methods, which has a very low participation of female economists; and another more gender-balanced club made up of Micro & Other fields.[12] Actually, the share of women in the AEA program displays a virtually constant difference of about 10 percentage points between these two broader research areas, which constitutes a considerable difference if we take into account the fact that the average female participation is less than 25 percent. Depending on the year, female authors represent between 16 and 20 percent of all authors in 'Macro, Finance and Methods', and between 27 and 30 percent of total authors in 'Micro & Other'.

In a recent paper, Chari and Goldsmith-Pinkham (2017) extract information from the NBER Summer Institute sessions, over the period 2001-2016, to calculate the share of women in finance, macro & international, and micro subfields. They also find a persistent gap between subfields, with figures that are closely aligned with those we offer here for those years.

[12]Methods is the smallest categories by number of authors, accounting for an average of 2.6 percent of total authors over the whole period.

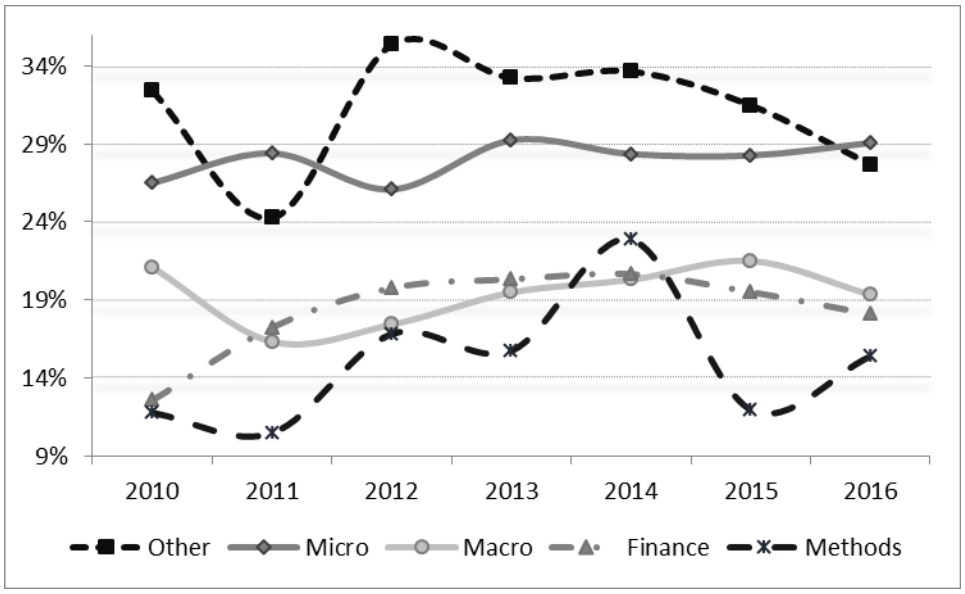


Figure 2 : *Share of women: research subfields (2010-2016)*

3.2 Abstract approach

Session JEL codes are quite general and may be imprecise indicators of what a paper is about; thus, the JEL approach might not be sufficiently fine-grained to capture the actual research topics covered in the meetings. To overcome this limitation, we dig into the paper abstracts and extract information using a machine learning technique. The website of the AEA meeting offers abstracts for the years 2014, 2015 and 2016. Using the Latent Dirichlet Allocation (LDA) algorithm for text analysis developed by Blei et al. (2003), we have classified the AEA abstracts by topics. The LDA algorithm offers a simple way of classifying large amounts of text (more than 5000 abstracts in this case)[13]. After some testing, we set a number of 21 topics with which to classify abstracts. We consider that the choice of 21 clouds satisfies a balance between the identification of underlying themes and the principle of parsimony.

For a more reader-friendly presentation of the words composing each of the themes obtained, we have constructed a word cloud for each theme. Figure 3 shows the three clouds with the highest and the three with the lowest share of women, averaging the three years for which we have information. The complete set of clouds, ordered according to the average share of women, can be found in Table 7 in Appendix A.

Topics related to gender, education and health appear as the most attractive to women, with the share of female authors ranging from 47.7 to 32.2 percent of the total. Conversely, at the bottom of the distribution of choices, we find themes linked with theoretical econo-

[13]Examples of the use of this algorithm in the economics literature can be found, for example in Hansen et al. (2017) and Muller et al (2017).

metrics, finance and macroeconomics, where women represent less than 16 percent of the authors. These results reinforce and complement our findings with the JEL approach. In both cases, they are fairly closely aligned with those obtained by Dolado et al. (2012) using a different methodology. These authors collected data from the personal websites of faculty members of the top-50 economics departments in 2005 as listed on Econphd.net. Using Econphd.net codes to classify research into different fields, they observed that topics such as wages and gender inequality, education, health and demographics were among the top female choices, with mathematical economics, fluctuations and business cycles at the other extreme.

Interestingly, our results indicate that themes conventionally aligned with macroeconomics are actually in the minority. In Table 7 (in Appendix A), we clearly identify only three out of the twenty-one clouds with macroeconomics (those assigned ranks 12, 16 and 21), to which we could add two finance-related clouds (ranks 11 and 19) plus one international economics cloud (rank 10) and one econometrics cloud (rank 20). Apart from three additional clouds that we find difficult to classify (ranks 6, 7 and 14), the remaining eleven could be classified as conventional microeconomic topics. This evidence suggests at least three ideas regarding the communication challenge that we described in the introduction. First, a fairly ample range of topics is covered by microeconomics, in contrast with the apparently prevalent idea among outsiders that economists essentially deal with macroeconomic issues and finance. Second, the fact that microeconomic topics are highly varied and correspond to a wide set of words complicates the task of conveying in a simple and straightforward way to those outside the academia that these topics are an essential part of economics. Finally, economics faces two main challenges when it comes to enhancing the visibility of female economists, namely, raising the profile of the first clouds in Figure 3 and boosting the proportion of women in the last clouds. It would be great if young students had a good understanding of these word clouds, their true contents, the human and social problems they deal with, and, most importantly from the perspective of this paper, were aware of the presence of women in these areas of research.

Having shown that there is considerable gender bias in the choice of research areas, which leads to an uneven gender composition across different economics subfields, we provide evidence in the following section that this gender-based inclination towards specific subfields appears earlier, at the undergraduate level.

4 Relative performance across fields: results with administrative data from the University of Valencia

In this section we use UV administrative records for the period 2010-2014, to estimate the relative academic performance of male and female students across subfields. First, we present

the data, then we estimate quantile regression (QR henceforth) for the students' grades in all courses classified into subfields (macroeconomics, microeconomics, finance, methods and other), and finally we estimate QR regressions using the data on optional courses correcting for selection into these courses by means of a semi-parametric selection model.

4.1 Data and quantile regression estimation.

The University of Valencia, with more than 50,000 students in 2017, is one of the largest public universities in Spain. Following a formal request for information, UV provided us with data for all students enrolled in economics at UV from 2010 to 2014. For all these students we received anonymized administrative records containing information on each student's entrance grade, the educational level and employment status of their father and mother, the student's employment status, courses that the student is or has been enrolled in over the sample years, exams taken each year (per course) and the grades achieved, as well as certain demographic characteristics such as gender and age.

In our data sample, each observation refers to one student in a given course and degree year. Each student contributes the same number of observations to our estimation sample as courses that the student is or has been enrolled in over the sample years. That is, we have not only summary information of a student's performance in a given academic year, but also their performance in all the courses they have taken over the sample period.

Figure 4 displays the trends in the percentage of students enrolled in the economics degree from 2005 to 2017 (though our estimation sample stops at 2014). The figure shows that, although the relative percentages of men and women were fairly well balanced in the years 2005 to 2007, the share of women has remained below that for men in the years since. Furthermore, and in line with what seems to be the general trend worldwide, in recent years the share of women has started to fall further and further below that of men. In the last year female students account for only 36 percent of the total.

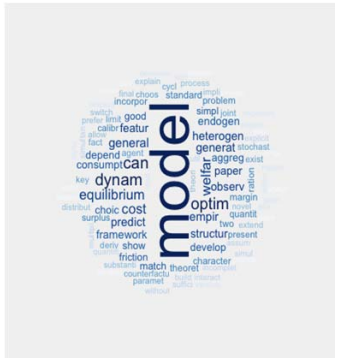
Highest Clouds			
Share of women	47.7%	35.1%	32.2%
No. authors	197	155	169
Lowest Clouds			
Share of women	15.8%	14.5%	11.3%
No. authors	249	123	102

Figure 3 : *Clouds of words from AEA papers' abstracts (2014-2016)*

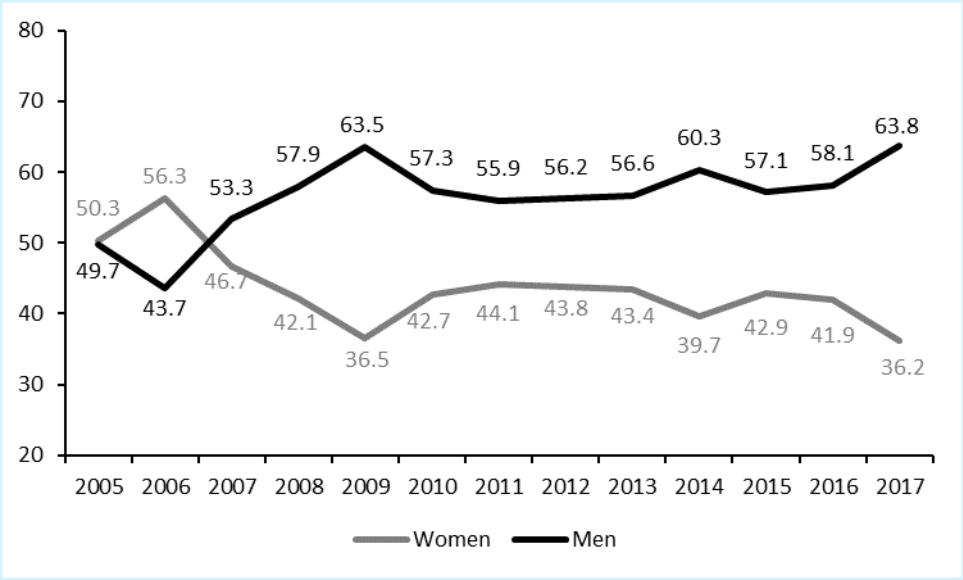


Figure 4 : *Percentages of male and female students in Economics (UV).*

The economics major at UV is organized around both compulsory and optional courses. Our sample data comprises both students entering under the Bologna Program launched in 2009 and students who had entered before that year and were still studying in years 2010 to 2014. Under the Bologna Program, students need to pass 240 ECTS credits over four academic years, with an average load of 60 credits per year. Courses can have different credit loads, with the most typical being 4.5, 6 or 9 credits per course. Courses are classified as core, compulsory or optional, and are spread out over the different years of the degrees. Optional courses are taken in the last year of the degree and represent 20 percent of the total amount of credits in economics. Under the former program, the 2000 Program, students had to pass 300 credits over five years organized in a three-year first cycle (180 credits) and a two-year second cycle (120 credits). Out of the total number of credits, 30 credits corresponded to optional courses taken over the three years of the first cycle, and 54 credits over the two years of the second cycle.

Similarly to what we did in the previous section, we classify all the courses into five subfields: macroeconomics, microeconomics, finance, methods and other. The courses included in each subfield are listed in Table 8 in Appendix B. Each student enters the estimation regression as many times as courses they have taken in the corresponding subfield, and we only consider the final grade when the exam was passed, that is, only one observation per student and course is used.

For each subfield we estimate a quantile regression to allow the results to differ along the grades distribution. The QR, introduced by Koenker and Basset (1978), estimates the quantiles of the conditional distribution of the response variable as functions of observed

covariates. Our QR regression for each subfield can be written as follows:

$$Q_q(G_{imt}) = \alpha(q) + \beta(q)W_i + \gamma(q)\mathbf{x}_{it} + m_t(q) + \tau(q) + Q_q(u_{imt}) \quad (1)$$

where Q_q denotes the conditional *qth* quantile; subscripts i, m and t denote the student, the course, and the year, respectively; W_i is a gender dummy that takes the value 1 if the student is a women and 0 if he is a man; G_{imt} denotes the grade obtained by student i in each course m and in a given year t, the course belonging to the corresponding subfield for which we estimate the equation; $\mathbf{x}_{it}$ denotes a vector of characteristics of the student and other socio-economic controls relating to their parents, some of which may vary over time. We pool all the individual observations corresponding to courses within a same subfield across our sample years and include a set of year dummies to control for the year effects $\tau(q)$. The term $m_t(q)$ refers to course-effects that we allow to vary over years: specificities of a given course in a given year, such as a change in the teacher or the difficulty of the final exam (such course-year effects are captured in the form of the average grade of all students in the course in the year). Finally, u_{imt} is the error term of the equation.

Among the variables included in $\mathbf{x}_{it}$ that refer to the student, we pay special attention to the score obtained by the student in the official examinations taken to gain entrance to the university (we call this the AU-score throughout the rest of the paper.[14] The AU-score is considered as a measure of the initial or innate academic ability of the student, and acts as a student-specific effect in the estimation. Other student-level controls included in $\mathbf{x}_i$, some of which may vary over time, refer to age, whether they are full-time students, whether they live in the family home during the academic year and whether students come to the university from standard high schools (current high school programs), from vocational (or professional) high schools, from older high school programs or by transferring from previous university degrees. The rest of the socio-economic controls refer to the educational and economic situation of the student's parents. These are binary variables taking the value 1 whenever the parents have a given level of formal education or have a given employment status, and 0 otherwise. The information is given for the student's father and mother separately.

The QR quantiles considered are the 0.25, 0.50 and 0.75 quantiles.[15] The interpretation of coefficient estimates is essentially the same as in the median regression with the following

[14]The AU-score is obtained by averaging the student's grades achieved at high school and the grade obtained in a specific examination compulsory for enrollment in public universities in Spain. Every year, faculties establish the minimum AU-score for enrollment in a given degree.

[15]In particular, we have estimated simultaneous quantile regression using the command *sqreg* in *Stata*. It produces the same coefficients as standard quantile regression for each quantile but *sqreg* obtains an estimate of the VCE via bootstrapping. QR offers two advantages over the option of splitting the sample: i) it uses the entire sample information and thus does not involve a loss of degrees of freedom; and ii) it avoids the sample truncation problem which arises by splitting the data according to observed values of grades.

difference: in the median regression the constant is the median of the sample, while in the QR the constant in, say, the 0.25 quantile is the 0.25th percentile for the sample.

4.2 Results.

Compulsory Courses

We distinguish in our estimation between compulsory and optional courses. First, Table 1 displays the results of the QR for grades obtained by students in the compulsory courses of the different subfields. The introduction of the AU-score is important in our setting, since, if the composition of the sample was biased towards, say, women having higher/lower ability than men, the observed advantage/disadvantage of women could be attributed to their different ability. In Table 1 the following results stand out. In macroeconomics and finance, the differential in women's grades in macroeconomics is negative although the estimates are not, in general, statistically significant (an exception is the case of the 0.50 quantile for finance, where the effect is significant at the 10 percent level). However, in microeconomics, female students exhibit a positive differential with respect to their male coursemates that becomes larger and significant in the 0.50 and 0.75 quantiles. Likewise, in methods, women clearly show positive differentials in all three quantiles, particularly in the 0.50 and 0.75 quantiles.

Notice that, in these upper quantiles, a difference of 0.2 points (microeconomics) or 0.15 points (methods) in the final grade may determine which students get a final Distinction in the course; thus, for top students, such a difference may matter quite a lot.[16]

Optional Courses

Next, we focus on optional courses in the different subfields. In these courses, the academic content is more specific than in compulsory or core courses, the groups are smaller and the teachers are more specialized in the course. Also, in optional courses, students are more likely to self-select according to their preferences. It may well be the case that the differences in women's and men's relative performance are more clearly revealed in these courses.

In this case, we offer our QR results with and without control for selection. We consider the convenience of accounting for selection in this case for a number of reasons. On the one hand, students may prefer some subfields over others just because they find them easier to understand or simply more interesting. In turn, a variety of factors may influence why

[16] According to the rules of the Faculty of Economics at UV, only 1 in every 20 students enrolled in a group can be awarded a Distinction. Students who are awarded a Distinction are then exempted from paying for one of the courses in the following academic year.

Table 1 : Quantile regression of women and men grades across subfields (Compulsory Courses)

	(1) MACRO	(2) MICRO	(3) FINANCE	(4) METHODS	(5) OTHER
			quantile 25		
Women	-0.010	0.014	-0.020	0.057**	0.026
	(0.024)	(0.018)	(0.014)	(0.024)	(0.016)
AU-score	0.138***	0.143***	0.109***	0.137***	0.150***
	(0.014)	(0.011)	(0.016)	(0.011)	(0.006)
Constant	4.456***	4.410***	4.851***	4.642***	4.017***
	(0.206)	(0.128)	(0.141)	(0.197)	(0.089)
			quantile 50		
Women	-0.023	0.104*	-0.089*	0.186***	0.037
	(0.051)	(0.055)	(0.050)	(0.049)	(0.028)
AU-score	0.259***	0.232***	0.263***	0.260***	0.231***
	(0.016)	(0.020)	(0.020)	(0.020)	(0.009)
Constant	4.744***	4.938***	4.476***	5.187***	4.617***
	(0.308)	(0.229)	(0.407)	(0.302)	(0.145)
			quantile 75		
Women	-0.088	0.224***	-0.080	0.154***	0.069**
	(0.066)	(0.058)	(0.073)	(0.058)	(0.034)
AU-score	0.325***	0.277***	0.322***	0.347***	0.256***
	(0.021)	(0.020)	(0.028)	(0.019)	(0.010)
Constant	5.556***	6.781***	5.219***	6.746***	6.047***
	(0.433)	(0.424)	(0.465)	(0.269)	(0.183)
N Obs.	4313	5469	3459	6135	16125

Bootstrapped errors in parentheses (500 replications). * p<0.10 ** p<0.05 *** p<0.01. All regressions include course-year effects, year dummies and the following set of socio-economic controls: student's age, full-time student, living at home, father with primary education, mother with primary education, father with university education and mother with university education. AU-score: student's score to access university (averaged grades from high school and the university access exam).

a student finds a given subfield more or less interesting. For example, students' attention may be drawn to a subfield because it deals with topics that seem particularly appealing, or just because it is perceived as an area of study that will facilitate their future job search. Students may also choose optional courses on the basis of their previous results in the most closely related compulsory courses. Of course, since there is a limited number of optional courses on offer within a degree, they may also choose some courses through a process of elimination.

On the other hand, the above-mentioned considerations may differ among subfields or by gender. In fact, it is often argued that women tend to make choices more based on their performance.[17] If this were the case, the way women and men self-select could prevent us from observing their underlying capacity or tendency to obtain different grades. If women dislike, say, finance and only choose it when they believe they may well get good marks in it, whereas men tend to select into finance regardless of their prospective performance, then the composition of the groups in finance could result in very similar grades even if men were, say, naturally better at this subject. What we are in fact aiming to capture is the extent to which women and men exhibit different relative performance that goes beyond such determinants of self-selection into specific subfields.

We apply Buchinsky's (1998) method to correct for the effect of selection in QR. Basically, the method extends Heckman's (1979) classic model to allow for non-normality. In QR, since we aim to correct selection in various quantiles, we cannot restrict the analysis to a symmetric normal distribution. The Buchinsky method consists in first estimating a univariate binary-choice model through a semi- or non-parametric estimator; in a second stage, the QR for grades is estimated adding a correction term derived from the first stage. We leave for Appendix C the detailed description of the method, and the particular way in which we apply the selection correction.[18]

Table 2 presents the QR results for grades in the optional courses across subfields, with and without correction for selection.[19] First, the results confirm the previous finding that

[17]Rask and Tiefenthaler (2008) provide evidence suggesting that women demonstrate higher sensitivity to previous grades in selecting further economics courses.

[18]Our semi non-parametric estimator, SNP henceforth, is that proposed by Gallant and Nychka (1987), which has the advantage over other semi-parametric estimators that it does not involve kernel smoothing. This SNP estimation reveals some interesting findings (displayed in Appendix C). i) Past performance in micro seems to affect the choice of micro subjects differently in men and women; in particular, women with good marks in micro are far less likely to choose macro courses than men while they are much more prone than men to choose micro-related optional courses in that case. ii) It seems that men are not driven by their past marks in finance to enroll in finance-related courses, whereas women are more likely to enroll in finance courses if they previously achieved good grades in finance. iii) Higher ability students, as proxied by the AU-score, tend to choose macro and methods to a greater extent, while they seem to be less attracted by finance. iv) Women have lower total estimated probabilities of entering macro, and finance, as compared to micro and other courses.

[19]The correction for selection is statistically significant in all subfields, and the choice of the order of the polynomial on the index $(X_i'\beta_s)$ is made on the basis of how many powers are significant in the OLS

women perform better than men in micro while performing relatively worse in macro-related courses. The correction for selection has two effects on these results: the magnitude of the estimated coefficient for the women dummy is larger in both the micro and the macro equations, and the significance of the estimates increases, at least in the 0.50 and 0.75 quantiles. In microeconomics, women outperform men in all three quantiles, while the differentials are somewhat greater in the upper quantile, where the difference is near 0.4 points. In macroeconomics, women (when the covariates have a value of 0) obtain average grades that are more than 0.5 points below the grades obtained by men in the upper quantiles, while the estimates in the 0.25 and 0.50 quantiles, although also negative, are not statistically significant. It is important to emphasize here that these are not differences that teachers could easily observe in their lists of grades. Instead, these are the differences that would show up if the groups of students were composed of students that are broadly comparable with respect to a list of characteristics, including their innate ability and the reasons why they enroll in these subfields.

The fact that the differences are more evident in the upper part of the grades distribution makes the obtained results even more relevant. The upper tail of the distribution of grades is where the level of competitiveness between students is expected to be the highest, and thus, where the natural differences, in terms of ability and preferences between men and women, if they exist, will show up more clearly. In addition, a positive/negative difference in this upper tail gives the individual the self-confidence that he or she is notably better at some courses than at others, what may condition male and female graduates' choices of future occupations. It can also be the case that employers seek to hire or select graduates drawing from this upper part of the distribution. By way of example, students planning to follow a postgraduate course or to undertake a PhD may be selected by supervisors on the basis of their marks in specific subjects. The observation that there are more men in macro-related occupations and more women in micro-related ones could then be a consequence of these results.

Preferences or ability?

Faced with these results, the question that naturally arises is what might be behind these differences. Is it a matter of preferences? Or, alternatively, is it a matter of ability, some element not totally captured by the AU-score? Given that macro- and microeconomics-related courses at the degree level may imply different degrees of analytical skills, and given that the results in Table 1 suggest that women perform relatively better in methods, we repeat our main QR results partialling out the effect of the student's average grade in these instrumental courses.[20]

regression on the median (see Appendix C).

[20]Ideally, we would like to use some variable to account for students' innate or initial mathematical ability

In Table 3 we repeat the QR results for macro, micro and finance, both for compulsory and optional courses, after controlling for the student's average performance in methods. All grades obtained by the student in compulsory and optional methods-related courses are averaged as a proxy for the 'technical' ability of the student. A first observation is that the higher the average performance in methods, the higher the grade in all courses; moreover, the effects are greater in macro and micro than in finance. As compared with previous results displayed in Tables 1 and 2, we observe that the magnitude of the coefficient for the women dummy is, in general, adjusted downwards, thus suggesting that part of the previously estimated effect might be positively correlated with the students' analytical ability. Accordingly, the positive differential in favor of women in micro is slightly attenuated, while the negative results for macroeconomics are reinforced quantitatively, at least for compulsory courses and for optional ones in the 0.50 quantile.

We conclude from these results that even if differences in analytical ability might play some role, they do not suffice to explain the estimated differences in grades. Hence, preferences arise as the most satisfactory explanation of the observed differences in performance across subfields. In the next section we examine what students' self-statements about macro- and micro-related courses reveal as regards their preferences and interests in these two categories.

5 Survey given to economics students: opinions and beliefs, and their impact on performance.

In this section, we present the results from the self-statement questionnaire given to students of economics enrolled at UV in February of academic year 2017-2018. First we describe the survey design and present the main information at a descriptive level, and then we conduct OLS and IV estimation of the determinants of students' self-reported interests and performance in macro and microeconomics courses. In this latter case, we also conduct an estimation of peer effects to highlight the social interaction of male and female students in macro and microeconomics classes.

5.1 Survey design and self-reported statements about subfields and the economics profession.

Intrigued by our results revealing the different relative performance of male and female students in macro and micro courses, we surveyed students of economics to inquire into their opinions about the two subfields of macro and microeconomics and to gather information on their beliefs about the economics profession. An anonymous questionnaire was given to

before entering the university. Unfortunately, the AU-scores do not allow us to separate mathematics marks from marks in other subjects such as languages or history.

Table 2 : Quantile regression for grades in optional courses across subfields (without and with selection correction)

	(1) MACRO No corr.	(2) MACRO Select. corr.	(3) MICRO No corr.	(4) MICRO Select. corr.	(5) FINANCE No corr.	(6) FINANCE Select. corr.	(7) METHODS No corr.	(8) METHODS Select. corr.	(9) OTHER No corr.	(10) OTHER Select. corr.
					quantile 25					
Women	-0.347	-0.277	0.196**	0.258**	0.162	0.186	0.050	0.005	0.000	-0.001
	(0.338)	(0.284)	(0.097)	(0.107)	(0.157)	(0.168)	(0.116)	(0.125)	(0.000)	(0.002)
AU-score	0.151	-0.024	0.162***	0.069**	0.186***	0.203***	0.141***	0.158***	-0.000	0.003
	(0.197)	(0.191)	(0.036)	(0.034)	(0.043)	(0.045)	(0.048)	(0.053)	(0.000)	(0.004)
Constant	5.034***	5.754***	3.708***	4.319***	5.025***	5.071***	3.179***	3.655***	4.747***	4.596***
	(1.716)	(1.699)	(0.525)	(0.475)	(0.879)	(0.944)	(0.739)	(0.814)	(0.017)	(0.073)
					quantile 50					
Women	-0.454	-0.534	0.210*	0.251**	0.033	0.064	0.071	-0.128	-0.030*	-0.020***
	(0.296)	(0.329)	(0.118)	(0.110)	(0.144)	(0.140)	(0.141)	(0.154)	(0.016)	(0.007)
AU-score	0.218*	0.125	0.329***	0.216***	0.222***	0.228***	0.226***	0.220***	0.185***	0.062***
	(0.132)	(0.171)	(0.032)	(0.043)	(0.038)	(0.033)	(0.066)	(0.062)	(0.012)	(0.015)
Constant	7.063***	6.557***	4.671***	4.781***	5.711***	5.267***	4.299***	5.069***	4.372***	4.800***
	(1.445)	(1.742)	(0.644)	(0.687)	(0.810)	(0.717)	(0.924)	(0.866)	(0.207)	(0.108)
					quantile 75					
Women	-0.508*	-0.547**	0.274**	0.376***	0.165	0.063	0.057	-0.251	-0.065	-0.021
	(0.302)	(0.270)	(0.114)	(0.111)	(0.166)	(0.147)	(0.159)	(0.178)	(0.045)	(0.014)
AU-score	0.232**	0.063	0.351***	0.260***	0.144***	0.135***	0.218***	0.253***	0.206***	0.120***
	(0.101)	(0.119)	(0.032)	(0.039)	(0.049)	(0.045)	(0.055)	(0.060)	(0.012)	(0.014)
Constant	8.474***	7.630***	5.610***	5.970***	6.992***	6.628***	6.738***	6.474***	4.739***	5.876***
	(1.484)	(1.351)	(0.610)	(0.597)	(0.965)	(0.928)	(1.005)	(0.884)	(0.264)	(0.162)
N Obs.	256	256	1624	1624	755	755	1062	1062	12734	12734

Bootstrapped errors in parentheses (500 replications). * p<0.10 ** p<0.05 *** p<0.01. All regressions include course-year effects, year dummies and the following set of socio-economic controls: student's age, full-time student, living at home, father with primary education, mother with primary education, father with university education and mother with university education. AU-score: student's score to access university (averaged grades from high school and the university access exam).

Table 3 : QR for grades across subfields: robustness to student's grades in Methods

	COMPULSORY COURSES			OPTIONAL COURSES (Sel. corr.)		
	(1) MACRO	(2) MICRO	(3) FINANCE	(4) MACRO	(5) MICRO	(6) FINANCE
			quantile 25			
Women	-0.018	0.012	-0.023	-0.359	0.229**	0.212
	(0.030)	(0.021)	(0.022)	(0.237)	(0.092)	(0.151)
AU-score	0.065***	0.076***	0.058***	-0.050	0.030	0.161***
	(0.014)	(0.010)	(0.012)	(0.141)	(0.033)	(0.051)
Methods' grade	0.718***	0.674***	0.466***	1.786***	1.090***	0.537***
	(0.068)	(0.044)	(0.060)	(0.473)	(0.228)	(0.206)
Constant	4.444***	4.292***	4.737***	4.512***	4.173***	3.653***
	(0.194)	(0.117)	(0.189)	(1.389)	(0.548)	(0.912)
			quantile 50			
Women	-0.008	0.035	-0.073	-0.594**	0.267**	0.059
	(0.049)	(0.053)	(0.055)	(0.300)	(0.108)	(0.115)
AU-score	0.134***	0.102***	0.159***	-0.060	0.086**	0.206***
	(0.019)	(0.021)	(0.024)	(0.134)	(0.039)	(0.033)
Methods' grade	1.113***	1.141***	0.994***	1.428***	1.501***	0.545**
	(0.091)	(0.081)	(0.107)	(0.471)	(0.234)	(0.212)
Constant	5.345***	5.335***	4.550***	6.372***	4.410***	5.096***
	(0.271)	(0.263)	(0.354)	(1.299)	(0.638)	(0.697)
			quantile 75			
Women	-0.147**	0.185***	-0.168**	-0.424**	0.208**	0.048
	(0.070)	(0.062)	(0.079)	(0.210)	(0.106)	(0.135)
AU-score	0.171***	0.133***	0.194***	-0.095	0.209***	0.118***
	(0.025)	(0.026)	(0.031)	(0.138)	(0.042)	(0.041)
Methods' grade	1.212***	1.237***	1.162***	1.357**	1.320***	0.367
	(0.111)	(0.090)	(0.119)	(0.551)	(0.202)	(0.284)
Constant	5.919***	6.469***	5.374***	8.186***	6.025***	6.122***
	(0.526)	(0.391)	(0.538)	(1.412)	(0.624)	(0.893)
N Obs.	4313	5469	3459	256	1624	755

Bootstrapped errors in parentheses (500 replications). * p<0.10 ** p<0.05 *** p<0.01. All regressions include course-year effects, year dummies and the following set of socio-economic controls: student's age, full-time student, living at home, father with primary education, mother with primary education, father with university education and mother with university education. AU-score: student's score to access university (averaged grades from high school and the university access exam).

students in class time with the collaboration of the teachers of the compulsory courses of the first, second and third year of the economics degree. The response rate is the number of students attending classes the day the survey, which was not announced in advance, was given out (N=307).[21] Besides their opinions about macro, micro and the profession, they were also asked questions concerning themselves and their parents' education and jobs. More concretely, a first group of these questions relate to the students' gender, AU-scores, whether or not they had considered other alternatives to the economics major before entering the university, the type of high school they come from, as well as questions about their parents' socio-economic characteristics, coinciding with the categories in the administrative data we used in Section 4.

Students' opinions about macro vs. microeconomics.

To gather the students' opinions about macroeconomics and microeconomics, we asked them to declare which of these two subfields they found to be more intuitive, more aligned with social problems (we call this indicator *social concern*), technically more complex and more difficult to pass. Two additional questions in this group were of particular interest to us: first, whether students find either macro or micro a more interesting subfield than the other; and second, in which subfield they obtain better grades. They could also take a neutral position and answer 'both equally'.

In Table 4 we show the students responses. First, we present OLS results for several dependent variables constructed as binary indicators taking the value 1 if the student rates a given subfield (macro or micro) as more clearly intuitive, technically complex, difficult to pass, etc. The estimated constant in each case measures the percentage of male students that choose the characteristic denoted at the top (intuitive, complex, and so on) as more connected to macro or micro, while the estimated women difference adds to the constant to determine the percentage of women that choose that answer (those who take a neutral position make up the difference to 100 percent).

We can sum up our results as follows. The percentage of men that find macroeconomics more intuitive, more technically complex and more difficult to pass and the percentage of men that believe these characteristics to be more closely linked to micro are fairly similar (roughly 30 percent for both macro and micro). However, women significantly find macroeconomics more technically complex and more difficult to pass than men, while women are more likely than men to answer that micro is more intuitive than macroeconomics (around 32 percent of men as compared to nearly 46 percent of women). As regards which subfield is more closely aligned with important problems in society around 60 percent of our stu-

[21]The response rate is around 70 percent, a total of 9 groups were surveyed and the share of women is around 46 percent, somewhat greater than the actual percentage of enrolled women, suggesting that absenteeism among women might be lower.

dents rate macroeconomics clearly above microeconomics, and women do not differ from men in this statement (although the estimated differential is negative, it is not statistically significant in our regressions). In the case of the more general question of which subfield is found to be more interesting, it also seems that macro is preferred by both men and women, although in this case women exhibit a differential positive answer in micro and negative in macro with respect to men that is statistically significant. Finally, as regards their performance in the two subfields, a much higher percentage of men respond that they obtain better grades in macroeconomics (nearly 41 percent as compared to the 14 percent that claim to obtain better grades in micro). The percentages of women stating better grades in either micro or macro are quite similar (a total of 26 percent in macro and around 22 percent in micro) but relative to men, their self- stated performance is statistically worse in macro and better in micro.

This first set of results, thus, reveals that women are more likely than men to choose micro as a more interesting subfield, and the opposite holds for macro. This does not mean that women find macroeconomics less interesting than microeconomics; it is just that they find it less so than men do. For reasons we are not able to identify, women find micro less difficult and more intuitive than macro, but, interestingly enough and in spite of this, they report similar performance in micro and macroeconomics. Unlike women, men exhibit more skewed answers towards a relatively better performance in macro. Thus, according to their own statements, women do not exhibit a particular preference or an absolute advantage in microeconomics as compared to macro but, instead, they clearly hold a comparative advantage/disadvantage in micro/macro as compared to men.

In the second half of the table, we repeat these OLS results controlling for students' ability (AU-score) and the rest of the controls. We also include the academic year in which the student is currently studying. This conditional estimation basically confirms the differential effects exhibited by women with respect to men outlined above.[22]

Students' beliefs about the economics profession.

Next, to assess the students' beliefs about the economics profession, we presented them with a list of seven possible occupations of an economist. For each one, we asked students to report their opinion concerning two aspects: first, whether they personally associate more clearly each of these occupations either with macroeconomics or with microeconomics (or both equally); and second, whether each of these occupations is perceived as a male or a female job (or, alternatively, no gender bias). The seven occupations listed refer to jobs in: 1. A research center in Spain. 2. A research center abroad. 3. Banking and/ or Finance

[22]Here the constants no longer admit the same interpretation as in the upper half of the table. They are now the proportion of men that chose that option when all the rest of the covariates have a value of 0 (including age, for example, which cannot be zero).

Table 4 : STUDENTS' SELF-STATEMENTS ABOUT SUBJECTS IN MACRO AND MICRO "I FIND SUBJECTS IN THIS SUBFIELD MORE..."

	Intuitive	Technically complex	Difficult to pass	Better grades	Socially concerned	Interesting
MACRO						
Woman (difference)	0.003	0.163***	0.168***	-0.142***	-0.0229	-0.112**
	(0.053)	(0.055)	(0.054)	(0.053)	(0.056)	(0.037)
Constant (avg. men)	0.307***	0.337***	0.283***	0.410***	0.614***	0.626***
	(0.035)	(0.036)	(0.035)	(0.038)	(0.037)	(0.038)
MICRO						
Woman (difference)	0.132**	-0.115**	-0.118***	0.073*	0.059	0.065*
	(0.055)	(0.049)	(0.045)	(0.044)	(0.043)	(0.035)
Constant (avg. men)	0.325***	0.319***	0.259***	0.145***	0.145***	0.090**
	(0.036)	(0.036)	(0.034)	(0.027)	(0.027)	(0.027)
MACRO						
Woman (difference)	0.026	0.135**	0.130**	-0.111**	-0.034	-0.114**
	(0.052)	(0.056)	(0.055)	(0.050)	(0.057)	(0.056)
Constant	0.301	0.767**	0.627**	0.067	0.556*	0.548*
	(0.240)	(0.300)	(0.245)	(0.264)	(0.312)	(0.303)
AU-score	0.002	-0.008	0.004	0.0113	0.008	0.017
	(0.016)	(0.020)	(0.018)	(0.016)	(0.020)	(0.019)
Age	-0.007	-0.011	-0.013**	-0.001	-0.001	-0.011
	(0.006)	(0.007)	(0.005)	(0.007)	(0.009)	(0.009)
Year 2	0.331***	-0.256***	-0.318***	0.517***	-0.007	0.298***
	(0.069)	(0.069)	(0.060)	(0.064)	(0.074)	(0.067)
Year 3	0.117*	-0.099	-0.038	0.285***	0.042	0.187***
	(0.061)	(0.067)	(0.069)	(0.059)	(0.071)	(0.070)
F-test controls[a]	yes	yes	yes	yes	yes	yes
p-values	0.708	0.590	0.250	0.865	0.504	0.927
MICRO						
Woman	0.116**	-0.110**	-0.075*	0.068*	0.067	0.061*
	(0.055)	(0.049)	(0.042)	(0.035)	(0.044)	(0.035)
Constant	0.903***	-0.139	-0.038	0.745***	0.162	0.316
	(0.299)	(0.250)	(0.263)	(0.239)	(0.237)	(0.220)
AU-score	-0.034*	0.033**	-0.001	-0.038**	-0.008	-0.008
	(0.017)	(0.015)	(0.013)	(0.018)	(0.015)	(0.017)
Age	-0.004	-0.001	0.00774	-0.008*	0.00477	-0.004
	(0.009)	(0.007)	(0.009)	(0.004)	(0.007)	(0.003)
Year 2	-0.323***	0.257***	0.424***	-0.102*	-0.0397	-0.122***
	(0.066)	(0.066)	(0.064)	(0.052)	(0.057)	(0.041)
Year 3	-0.170**	0.206***	0.068	-0.034	-0.087*	-0.058
	(0.071)	(0.057)	(0.046)	(0.057)	(0.052)	(0.048)
F-test controls[a]	yes	yes	yes	yes	yes	yes
p-values	0.064	0.398	0.309	0.563	0.296	0.955
N Obs	307	307	307	307	307	307

Robust standard errors in parentheses. * p<0.10 ** p<0.05 *** p<0.01. [a] Controls: student's age, full-time student, living at home, father with primary education, mother with primary education, father with university education and mother with university education. AU-score: student's score to access university (averaged grades from high school and the university access exam).

sector. 4. PhD to become an academic. 5. Advisor in institutions (public or private). 6. Private sector. 7. Civil servant (through official examinations). Students were also allowed to specify any 'other' job they might contemplate. We believe that most would agree that the list itself is neither field-biased nor gender-biased.

The students' answers as regards their perception of the listed occupations reveal that their beliefs about the economics profession are affected by subfield and gender biases. These answers are shown in Figures 5 and 6. The first figure illustrates the percentage of students that associate each of the listed jobs either with macro or with microeconomics (or neutral). The data reveal, in our view, some clear mistaken beliefs held by students in this regard. For example, they associate jobs in research centers more with macro than with micro, and even more with macro than with 'both': around 52 percent and 64 percent, respectively, make such an association, against 10 percent and 3.6 percent that regard these two occupations as more connected with microeconomics. Becoming an advisor in institutions, no matter if they are public or private, is also prominently associated with macroeconomics (65 percent of students believe this, against 24 percent that associate it with both subfields, and 7.8 percent that link it to micro). The banking and finance sector is again more strongly associated with macroeconomics. The private sector is the only case where microeconomics clearly outnumbers macroeconomics (78 percent of respondents connect it more clearly to micro). Finally, micro is approximately equal to macroeconomics as a subfield that is specific to civil servant, that is, jobs in the public sector obtained by sitting official examinations, and finally, curiously enough too, microeconomics is somewhat more associated with following a PhD and becoming academics (18 percent for micro as compared to 12.7 percent for macro).[23]

At least three concerns arise from this information. First, the economics profession is viewed as dominated by macroeconomics in most of economists' more natural occupations. The many interesting fields where microeconomists work and their relevance in providing advice to public and private institutions do not seem to reach our students. Second, this imperfect information is predictably much more pronounced among young students before their choice of major. Those not particularly attracted by macroeconomics are very unlikely to feel attracted by economics, in part due to the lack of appropriate information as regards the many topics that microeconomics deals with. And, third, if microeconomics is particularly associated with only the private sector (probably viewed as merely dealing with firms' and consumers' decisions) it then follows that business majors offer a natural alternative for

[23]In general, no significant differences between women and men emerge in these answers, so we do not offer them separately by gender. The only statistically significant differences are that 4 percent (p-value=0.072) and 6 percent more women (p-value=0.027) find jobs in research centers abroad and jobs in finance and banking, respectively, to be more associated with women, and 12 percent more women (p-value=0.038) regard jobs in research centers abroad as more associated with macroeconomics.

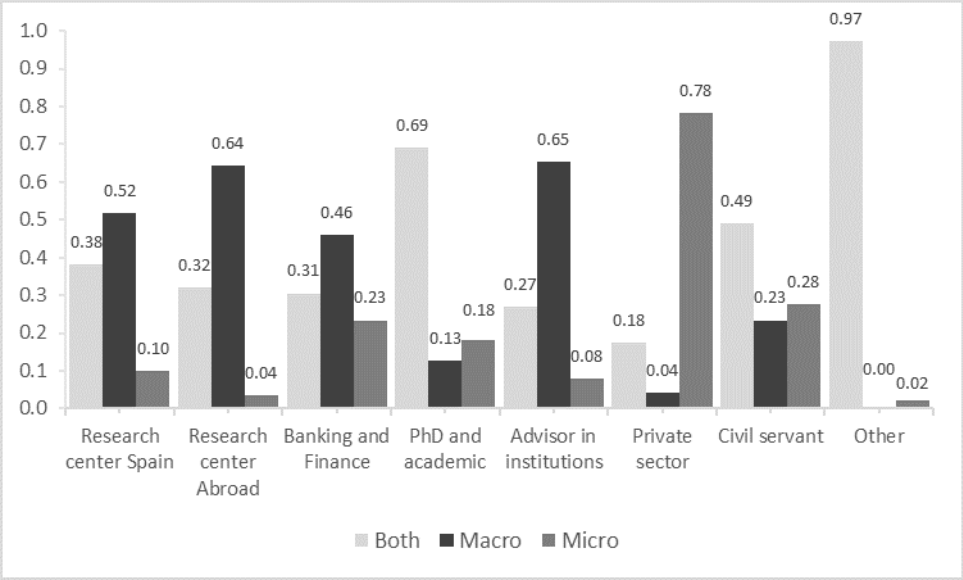


Figure 5 : *Students' association of jobs with MACRO vs MICRO*

those who think they will enjoy microeconomics more.[24]

Moving to Figure 6 we can observe the students' beliefs about the gender bias of the different types of occupations. A first observation to note here is that, in all cases, the highest percentage corresponds to students that find the job equally associated with women and men. In any case, when students side with any gender, it is more often a man than a woman in most of the occupations: in two out of the seven cases (finance and advisor to institutions) more than 40 percent of students find the job to be particularly associated with a male figure (vs. 6.5 percent associating it with women), and in three out of the remaining five jobs, around 20 percent of students provide the same answer (no more than 6 percent associate them with women in any case). In only two out of the seven occupations, women are slightly more salient: following a PhD and becoming an academic, and sitting official examinations to work in the public sector. These are the two jobs in the list that probably entail a lower perceived risk and greater stability, namely, an academic and a civil servant. We find it quite curious that the two occupations less clearly associated with macro are precisely those that are more closely associated with women.[25]

Thus, for many of our students, their overall perception of the profession seems to be

[24]At UV the percentage of women in business majors is slightly above 50 percent. Since we do not have information on students before they choose their majors nor on students that have decided not to enter economics, we cannot ascertain the extent to which these misperceptions about the profession may affect such decisions. Very likely, though, their understanding of the economics field is even worse in such cases.

[25]Social psychologists point out that individuals' assertions when asked about specific biases tend to hide the true ones. To overcome this recognized problem, psychologists carry out *Implicit Association Tests*, IAT, designed to detect the strength of a person's automatic association between objects or ideas (Greenwald et al., 1998).

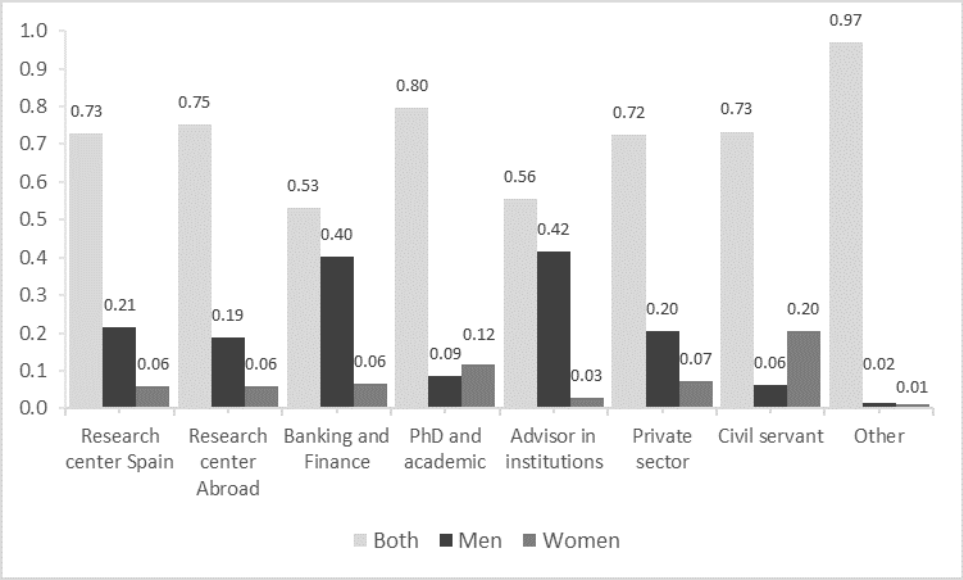


Figure 6 : *Students' association of jobs with MEN vs WOMEN*

one that offers greater opportunities for macroeconomists, particularly if they are men. It is difficult to see how such an environment could be congenial to women. Not only they do not belong to the dominant gender in some of the most visible occupations of the profession, but nor do they have a relative advantage in or preference for the dominant subfield, macroeconomics. The most disturbing fact, however, is that some of these established stereotypes respond to mistaken beliefs about the profession. It seems, for example, that the important role of microeconomics in providing advice to policymakers in the areas of education, health, energy, industry or labor is not clearly understood by our students. Policy advice in these areas is viewed as a task carried out by macroeconomists. The data also reveal a considerable misunderstanding of the job opportunities offered by the profession, as well as how the subfields they study in the degree translate into specific occupations in the labor market.

5.2 Students' perceptions and contagion effects on interests and grades.

5.2.1 Empirical equations and estimation procedure.

We now extend some of the baseline regressions offered in Table 4 by incorporating a full set of covariates. In particular, we focus on the indicator variables revealing which subfield students find more interesting and in which one they claim to obtain better grades. Our goal here is to find out, on the one hand, whether students' preferences for a particular subfield may affect performance, and, on the other hand, how interests and grades correlate with other students' statements regarding themselves as well as their beliefs about the profession.

In this section we also pay special attention to the likely contagion effects among students, both in their interests for the two subfields of macro and microeconomics and in the form of learning externalities in the classrooms.

We proceed as follows. For each one of the two subfields, we estimate a linear probability model (LPM henceforth) for the dependent binary indicator I_{if} equal to 1 if the *i-th* student states "*I find subfield f more interesting* (f=macro, micro)" and 0 if the student reports that the other subfield is more interesting or takes a neutral position. The estimation equation for each subfield takes the form:

$$\begin{aligned} I_i &= \alpha + \beta W_i + \delta \mathbf{z}_i^I + \gamma \mathbf{x}_i + \tau_t + \\ &\quad + \phi_W^{I^w} W_i \times PI_{j-i}^W + \phi_M^{I^w} M_i \times PI_{j-i}^W + \\ &\quad + \phi_W^{I^m} W_i \times PI_{j-i}^M + \phi_M^{I^m} M_i \times PI_{j-i}^M + u_i^I \end{aligned} \tag{2}$$

where W_i is the women dummy variable. The vector $\mathbf{z}_i^I$ contains the student's answers to other questions of the survey that we consider may correlate with the perceived interest in the subfield: a dummy variable that equals 1 for those who declare the subfield to be more closely linked to social problems; a dummy variable identifying those for whom economics was their first best option among all the university degrees (*initial vocation*); a quantitative variable counting how many jobs the student perceives as macro-oriented (we call this the "*Jobs-macro*" belief); a quantitative variable counting how many jobs the student perceives as macro-and-male-oriented (we call this the '*Jobs-macro-men*' belief); $\mathbf{x}_i$ denotes a vector of other characteristics of the student and family socio-economic controls, and τ_t stands for academic-year fixed effects (a set of dummy variables to account for the academic year the surveyed student is currently undertaking - first, second or third year of the degree). The rest of variables in equation 2 refer to our measures of male and female peers' interests that we construct as follows. For each student in each group we define the percentage of their female and male classmates who claim to find that particular subfield more interesting than the other (we label these peer measures PI_{j-i}^W and PI_{j-i}^M, respectively), and we cross them with the dummy variables for women and men, W_i and M_i, to allow the responses to peers to differ by gender. Finally, u_i^I is the error term of the equation. We estimate an equation like 2 separately for macro and for microeconomics.

We also define a similar model for the students' self-reported statements regarding their relative performance, where we consider, in turn, that the interest in that field may be among the explanatory factors. The estimation equation in this case can be written as:

$$
\begin{aligned}
G_i &= \alpha + \beta W_i + \lambda I_i + \delta \mathbf{z}_i^G + \gamma \mathbf{x}_i + \tau_t + \\
&\quad + \phi_W^{G^w} W_i \times PG_{j-i}^W + \phi_M^{G^w} M_i \times PG_{j-i}^W + \\
&\quad + \phi_W^{G^m} W_i \times PG_{j-i}^M + \phi_M^{G^m} M_i \times PG_{j-i}^M + u_i^G
\end{aligned} \tag{3}
$$

where G_i stands for the binary indicator taking the value 1 if the student states that their grades in macroeconomics (microeconomics) are better than in microeconomics (macroeconomics) or are the same. The equation for grades incorporates on the right-hand side the indicator of interest, I_i, among its explanatory factors. $\mathbf{z}_i^G$ denotes a vector containing the student's answers to other survey questions that we consider may correlate with their relative grades in the subfield, containing the same variables as $\mathbf{z}_i^I$ except for *initial vocation* and *social concern*.[26] Similarly to equation 2, we construct here the percentage of a student's female and male classmates who claim to obtain better grades in that subfield (PG_{j-i}^W and PG_{j-i}^M, respectively), and cross them with the gender dummies. We estimate an equation like 3 separately for macro and for microeconomics.

Identification of the causal effect of peers entails certain difficulties, endogeneity being a clear concern in this case. Among other reasons, the peer effect is endogenous because of an inherent reason: if peers' outcomes affect individual outcomes, individual outcomes will affect peers' outcomes, leading to simultaneity bias.[27] Here we follow the most widely-used approach to overcome the endogeneity problem, which consists in instrumenting the peer measures (e.g., Goux et al., 2007). In particular, we use as instruments two variables that are predetermined before students enter their respective groups and before their interaction in the classroom starts: the AU-scores and their self-reported statements about initial vocation. Following the standard practice, we aggregate the individual responses to these two variables,

[26] In exploratory work, initial vocation and social concern did not yield statistical significance in the grades equation, which indicates that, beyond their impact through interest, they do not directly affect relative grades. This means that these variables are naturally excluded from the equation for grades, which they only impact through their effect on the interest in the subfield. Hence, the two equations, interest and grades, can be thought of as the first and second stages, respectively, of an instrumental variables setting where the covariates excluded in the equation for grades can be used as instruments for interest to assess its exogeneity in the second stage. This allowed us to test for exogeneity and validity of instruments using initial vocation and social concern as instruments for interest in the equation for grades. These results allowed us to rule out such endogeneity concerns. These results are available upon request.

[27] Selection bias and other unobserved influences are two additional sources of endogeneity. The so-called *reflection problem* is one of the main sources of such endogeneity. It refers to the endogenous self-selection of individuals into their groups of influence, that is, individuals approach those whose unobservable characteristics are systematically related to theirs. An additional problem for the estimation of peer effects is that, in many educational empirical studies, data do not include class identifiers, which impedes the selection of a student's relevant peers. Our survey data allows us to identify the class, which permits the construction of the peer measures for each group of students spending class time together at university. The administrative data used in Section 4, for example, do not allow us to identify the class groups.

which then serve as IVs. Appendix D details the 2SLS procedure followed.[28]

5.2.2 Estimation results.

Relative interests.

Table 5 reports the estimation results for the determinants of interests and grades on the two subfields. In all cases, we interact the explanatory variables with the women dummy variable to allow the effects to differ by gender.

A first interesting result is that students entering economics with an initial vocation are around 25 percent more likely to answer that they find macroeconomics more interesting than micro. In micro, this variable correlates negatively with students' opinion that this subfield is more interesting, with the estimate being statistically significant at the 10 percent only in the case of women. We find this unequal correlation between initial vocation and interest in the two subfields very relevant because of the likely implications. On the one hand, the fact that both entry into economics as a vocation and interest in the subfield correlate positively with macro but negatively with micro may reflect the idea that students have a serious misunderstanding of what economics is about. On the other hand, these mistaken beliefs are biased against microeconomics and, thus, they may dissuade from enrolling those that could be more naturally interested in topics addressed by microeconomic theories and empirics, were they aware of them. In gender relative terms, as a higher proportion of women seem to be interested in micro, this could help explain the lower percentage of women that opt for economics as a vocation.[29]

The results also support the idea that, if a student views a given subfield as more aligned with social problems, this correlates positively with their statement that this subfield is more interesting than the other. However, while in macroeconomics the effect is common to both men and women (an estimated coefficient of 0.161), in the case of microeconomics the effect is appreciably higher (around 0.227) and significant only in the case of women. This indicates that, for women, the preference for microeconomics is fairly closely related to the degree to which this subfield is perceived to be connected to social problems.

[28]The use of LS to estimate in the second stage our binary variables of interests and grades yield average partial effects, APE, that prove to be virtually the same as the marginal effects on the average from an alternative Probit estimation. Given its greater flexibility and ease of interpretation, we apply LS to the mentioned two binary indicators.

[29]In a regression of the variable *initial vocation* on the indicator of macro as a more interesting subfield and on the number of occupations associated with macroeconomics (plus socio-economic controls), we found significant positive correlations in the first case (coefficient of 0.077 with p-value=0.031) and in the second case only for men (coefficient 0.158 with p-value=0.045). Thus, entering economics as a vocation is correlated with the interest in macroeconomics, and, in the case of men, also with the number of occupations associated with macro. The women dummy variable is negative in the regression, although not statistically significant. These results correspond to the sample of students in the first year, and are available on request.

An additional finding in Table 5 is the fact that, for women, believing a larger number of occupations to be linked to macroeconomics seems to convey the idea that macro is more interesting than micro, while this variable does not render statistical significance in the case of men. This latter finding highlights that female students tend to find macroeconomics more interesting as the number of occupations perceived as being held by macroeconomists rises; a practical view.

At the bottom part of columns 1 and 2 in Table 5 we display the results for the peers' variables.[30] These have been standardized for ease of interpretation. Thus, the estimated coefficients measure marginal probabilities of one standard deviation from the mean of the corresponding peer measure. In the case of macroeconomics, the results suggest that the larger the percentage of women in the classroom who show a clear interest in that subfield, the higher the probability that a female student in that classroom will claim to find it interesting too. The estimated coefficient suggests that if the percentage of women in a classroom with a marked interest in macroeconomics (as compared to micro or being neutral) increased by one standard deviation above its mean, the likelihood of a female student declaring that macroeconomics is more interesting than micro would increase by nearly 20 percent. Unlike women, men do not seem to be affected by their peers' interests in macroeconomics. The coefficient, albeit positive, is comparatively low and not significantly estimated. In microeconomics, however, the results suggest that men are positively influenced by their male peers (coefficient of 0.042 with p-value=0.019), and women by their female classmates (coefficient of 0.074 with p-value=0.040). Thus, in microeconomics the estimated effects are lower than in macro, and also lower for men than for women.[31]

These results for the effects of peers' interests in one student's own interests admit the following interpretations. First, the responsiveness of students to their peers' interests, when they exist, is linked to peers of the same gender. Women do not respond to the percentage of men that declare a greater interest in a given subfield, and nor do men to their female peers' interests.[32] Second, men are less responsive than women to their peers' interests. In macro, they are not significantly affected at all, and in microeconomics, though

[30] A priori, the signs of these peer or contagion effects are unknown to us. In principle, students will be more strongly influenced by those classmates with whom they most often spend time, exchange opinions or share information, who tend to be of the same gender. Conversely, students may be more influenced by those who are perceived as the best at a given subfield, regardless of gender. Perceptions, whether or not they are right, that men or women are relatively better at one subfield than the other may also condition which individuals play a leading role in a classroom and are more capable of generating contagion effects. Furthermore, the reaction to these effects may also be gender and subfield specific depending, for example, on the student's own independent interests or abilities. We are not aware of any previous work that deals with peer effects formulated in this way.

[31] At the bottom of the table, the F-statistics for joint significance of IVs in the first stage take considerably large values, thus ruling out concerns about the weakness of instruments. A full set of group dummies is also included in the estimation to control for other possible group-specific unobservables.

[32] The principle of *homophily* states that individuals prefer other people who are similar to them (McPherson et al., 2001), gender being one of the main aspects of such similarity.

the estimated effect is positive and statistically significant, its magnitude is the lowest among all the estimated impacts. The fact that men do not react to their peers' interest in macroeconomics may be due to their greater individual interest in macroeconomics, which probably exists independently of feedback from others. In micro, a subfield in which they are not on average particularly interested, a higher proportion of interested male classmates may attract their attention to this subfield. In other words, it seems that the lower the initial individual interest, the higher the potential for peers' influence to make a difference. This could also help explain the fact that the peer effect for women in macroeconomics is appreciably larger than in microeconomics.[33]

Summing up, vocational students tend to find macroeconomics more interesting than micro, the perception of a larger number of occupations for macro leads to a greater interest in macro on the part of women, and there seem to exist contagion or *social multiplier effects* (Glaeser et al., 2003) in the interests of female students towards macroeconomics. It is worth noting that, after discounting the observable factors, the women dummy variable still points out a lower relative interest of women in macro than men.

In the case of microeconomics nothing is left in the women dummy that indicates a greater preference for microeconomics on the part of women, beyond those who find it more socially concerned. This indicates that the higher preference of women for microeconomics as compared to men comes basically from those women who find microeconomics more socially concerned that macro. Interestingly enough, contagion effects are much less pronounced than in macro, even if also here women seem to be more sensitive to their mates than men are.

Relative performance.

Next we comment on the results for the estimation of students' self-reported statement that they perform relatively better in one of the two subfields. Now we add the variable that captures whether students assess a given subfield as the most interesting one, which we further interact with the women dummy.

Women tend to declare better grades in macroeconomics when they also declare to find it more interesting than micro (only when the women dummy is interacted with "finding macro more interesting" the effect is positive and significant). For the case of microeconomics, the variable indicating a higher interest in this subfield has a large and statistically significant effect, both for men and for women. Although the effect is not significantly different for women, since women report a higher interest in micro relative to men, this could help explain

[33]Our findings here line up, for example, with Huntington-Klein and Rose (2018), who find that women's preferences and decisions are affected by greater exposure to women peers, while peer gender has little effect on men. The authors argue that the results may depend on the context, with more potential for peer effects to have an impact when there are fewer women.

why women tend to outperform their male coursemates in microeconomics. That is, more women than men show interest in micro, and this turns out to be a significant explanation of their better grades in this subfield.

An interesting result refers to the relationship between the AU-score of the student and their reported relative performance. For macroeconomics in column 3, results show that men are more likely to report a better performance in macroeconomics the higher their AU-score is. Nevertheless, the interacted term of the AU-score with the women dummy is negative, significant and clearly larger that for men, which for women would yield a (smaller but) comparable effect in quantitative terms but with the opposite sign. This result shows that, unlike men, the higher women's AU-score is, the much less likely they are to claim a relatively better performance in macroeconomics. As a result, the relative negative performance of women captured by the isolated women dummy vanishes (the sign is reversed and it is not statistically significant).

Hence, female students of higher ability tend to report worse grades in macroeconomics. This result adds further support to our main finding in Section 4 using the administrative data, which reveals that the largest differences between men and women correspond to the upper tail of the grades distribution.

Finally, our results in column 3 indicate that the belief that the profession is male- and macro-dominated has a significant and appreciable impact on the probability of a woman declaring that she obtains worse grades in macro than in micro, while it does not influence in any way the statements offered by men. The coefficient estimate for women suggests that for each additional job perceived by a female student as male- and -macro- dominated, she is more than 10 percent more likely to claim that her best grades are not achieved in macroeconomics than a woman who does not hold such a view of the profession.

It is hard to make a categorical judgement as to the reasons why this might happen. The most straightforward explanation for such an effect could be that, as the perception of the economics profession being macro- and male-dominated increases, women's discomfort in the degree increases, and this becomes particularly noticeable in their performance in macroeconomics. It could be thought of as a response explained by some degree of stereotype threat: the more convinced a woman is that the profession is dominated by male macroeconomists, the greater her conviction that men outperform women in macroeconomics and, thus, the worse the performance of women in macroeconomics.[34] Alternatively, a relatively worse

[34]Stereotype threat refers to a situation where a person is concerned with a negative stereotype about their social group (Steele and Aronson, 1995). This concern may cause the person to perform worse when competitors are thought to be better than her/himself. Spencer et al. (1999) found that, unlike men, women's performance in math is negatively affected by the risk of being judged according to the negative stereotype that women have weaker math ability. They hypothesized that the apprehension it causes may disrupt women's math performance. In a recent paper, Iribarry and Rey-Biel (2017) find that women underperform men when the task in which they compete is perceived as favoring men and they are explicitly informed of the presence of a strong rival.

performance of women in macroeconomics as compared to micro caused by the macro/male-dominated profession belief could be reflected in women's self-selection into given subfields. That is, women might react by focusing their attention and academic effort on the subjects they find more congenial to them and where they expect lower competition with men, a type of female reaction with broad support in the literature (see, e.g., Geenezy and Rustichini, 2004, -and references therein- where the authors provide experimental evidence that women dislike competition when it is against men).

Finally, we wonder to what extent the gendered differences found in performance may condition the ambient lived by students in classes, and whether this might act as a mechanism that reinforces the differences. Peer effects in grades may come from two main sources. The first is through joint study and accomplishment of tasks by sharing doubts and engaging in discussion, as well as other forms of learning spillovers. The second comes from competition among classmates' peers, which may turn out to have positive or negative effects, depending on each individual's reaction to competitive settings.

Our estimation results for peer effects do not show negative effects in any case. On the contrary, all significant estimates are positive, and they exhibit a gender and subject bias different to those obtained for interests. In particular, statistically significant peer effects in macroeconomics originate only from male students, but affect both male and female classmates (with similar estimated coefficients of 0.156 and 0.149, respectively). Neither men nor women seem to be affected by their female peers' performance in macroeconomics. Quite interestingly, though, in microeconomics the patterns differ. In this case, a male student is more likely to indicate a better relative performance in microeconomics as the percentage of male classmates with better grades in micro increases. The same applies to women with respect to their female peers.

These results for peer effects on grades might appear puzzling at first sight. In brief, men seem to collaborate/compete only with their male peers, in both subfields, while women collaborate/compete with classmates of different genders depending on the subfield, namely, male peers in macro and female peers in micro. An initial interpretation of these results could be that both men and women may hold their own ideas as regards which students are the top performers in the classrooms. In this sense, we might argue that men hold gender-biased opinions as regards performance; more specifically, male-biased perceptions of the classmates to look up to. Conversely, for women, the *a priori* could be based on precisely what our results in Section 4 have indicated, that is, men tend to outperform women in macro, while women outperform men in microeconomics. To the extent that women perceive a different relative performance of their male and female peers in their classes, they may become convinced that men are the leaders in macro while women are so in microeconomics. Thus, women would have what we could refer to as a performance-based

or statistical bias.[35] This first interpretation would line up with the results in Grunspan et al. (2016). The authors, using classroom data from an introductory undergraduate biology class at a large American university, find a strong male bias among males when assessing their perceptions about the performance of their female classmates, whom they clearly underrate. In contrast, females seem to rate their peers on the basis of their performance, regardless of gender.

There may be alternative explanations for these contagion effects on grades. Studies have shown that men tend to participate more actively in the classroom than women (Lundeberg and Punćcohaŕ, 1994), and it would be expected that more so the greater their interest in the subject matter. According to this view, if men contribute with more comments, questions and opinions in macroeconomics classes, in which they tend to be more interested, they probably gain greater recognition from both their male and female classmates. This, in turn, generates a situation whereby students in the classroom, no matter if they are men or women, are more likely to share tasks, ask for help or even compete with men. In microeconomics, the argument would follow that women fail to convey such an image, and each student collaborates then with peers of the same gender.

Summing up, our results with the survey data confirm that women underperformance in macroeconomics as compared to men occurs mainly in the upper part of the academic ability distribution, that finding macroeconomics interesting is a necessary condition for women to report better grades in this subfield, and that the fact of observing a larger amount of jobs associated to men and macro may impact negatively the performance of women in this subfield. In microeconomics, after discounting the effect of interest, the women dummy still points out that female students are more likely than men to state that micro is the subfield where they relatively obtain better grades.

Finally, and as a likely reflection of the outstanding role of men in macroeconomics, they are the only ones to influence the performance of their mates in macroeconomics classes. Women reporting relatively better grades in macro do not have any influential effect on the grades of their peers, and in microeconomics they only influence their female classmates.

6 Summary and conclusions

Academics in economics have long observed that, at the research level, women tend to self-select differently across subfields in our discipline. In particular, women are relatively more present in some of the research topics that we would classify under microeconomics. In this paper, we have provided new evidence of this fact using web scraping and machine learning techniques to gather information from the papers presented at the annual AEA meeting

[35] Statistical bias or inference (Phelps, 1972) refers to the case whereby people use group statistics as a proxy of the unknown, or imperfect, information about individuals.

Table 5 : Student's self-statements:
"I find subjects in this subfield more interesting" (Interest)
"I obtain better grades in subjects of this subfield" (Grades)

	INTEREST		GRADES	
	MACRO	MICRO	MACRO	MICRO
	(1)	(2)	(3)	(4)
Women	-0.993**	-0.243	0.114	0.992*
	(0.446)	(0.595)	(0.728)	(0.599)
Interesting			0.055	0.284*
			(0.076)	(0.153)
W*Interesting			0.180*	-0.116
			(0.098)	(0.181)
AU-score	0.021	-0.025	0.045*	-0.015
	(0.025)	(0.021)	(0.023)	(0.025)
W*AU-score	0.029	0.041	-0.079**	-0.036
	(0.040)	(0.031)	(0.036)	(0.037)
Initial Vocation	0.249**	-0.011		
	(0.091)	(0.053)		
W*Initial Vocation	-0.034	-0.132*		
	(0.143)	(0.075)		
Social concern	0.161**	0.025		
	(0.077)	(0.065)		
W*Social concern	-0.090	0.227*		
	(0.115)	(0.121)		
Jobs Men	0.018	0.026	-0.001	0.025
	(0.038)	(0.026)	(0.040)	(0.026)
W*Jobs Men	-0.011	0.005	0.018	-0.020
	(0.042)	(0.044)	(0.036)	(0.034)
Jobs Macro	-0.004	0.024	0.033	-0.017
	(0.031)	(0.016)	(0.029)	(0.014)
W*Jobs Macro	0.080**	-0.016	0.028	0.008
	(0.032)	(0.029)	(0.033)	(0.035)
Jobs Macro-Men	0.021	-0.033	-0.073	0.053
	(0.063)	(0.044)	(0.061)	(0.049)
W*Jobs Macro-Men	0.024	0.003	-0.118*	0.014
	(0.073)	(0.073)	(0.066)	(0.065)
Peers' Effects				
Men*Male Peers'	0.057	0.042**	0.156***	0.063*
	(0.087)	(0.019)	(0.057)	(0.033)
Women*Male Peers'	0.037	0.011	0.149**	0.010
	(0.069)	(0.020)	(0.073)	(0.044)
Men*Female Peers'	0.037	0.021	-0.010	-0.027
	(0.064)	(0.031)	(0.069)	(0.039)
Women*Female Peers'	0.198***	0.074*	-0.009	0.088**
	(0.070)	(0.040)	(0.090)	(0.045)
Constant	0.278	0.109	-0.302	0.130
	(0.587)	(0.231)	(0.370)	(0.332)
IVs for Peers	Peers' Vocation Peers' AU-scores	Peers' Vocation Peers' AU-scores	Peers' Vocation Peers' AU-scores	Peers' Vocation Peers' AU-scores
First stage F-test (IVs for male peers)	1336.57	319.18	321.17	441.49
First stage F-test (IVs for female peers)	613.58	363.09	103.27	195.71
N Obs	307	307	307	307
Adj. R^2	0.143	0.063	0.244	0.085

Bootstrapped standard errors in parentheses. * p<0.10 ** p<0.05 *** p<0.01. Controls tested and excluded because of lack of joint statistical significance: student's age, full-time student, living at home, father with primary education, mother with primary education, father with university education and mother with university education, plus year dummies. AU-score: student's score to access university (averaged grades from high school and the university access exam).

from 2010 to 2016.

However, to our knowledge, thus far there has been no evidence as to when these differences might arise. In this paper we have shown that they seem to start as early as at the undergraduate level, if not sooner. Using administrative data for economics students at a large public university in Spain we find that, relative to one other, male and female students perform differently in macroeconomics and microeconomics-related courses, more appreciably in the upper quantiles of the grades distribution.

The observed different performance of students by gender and subfield could be broadly attributed, in principle, to two potential factors. The first is possible differences in mathematical and analytical ability by gender. Our results after controlling for the students' overall academic ability and their specific methodological ability rule out this possibility. The second is different preferences for subfields by gender. To gather some information as regards this latter possibility, we surveyed students of economics enrolled at UV in the academic year 2017-2018. Some of their responses are striking, though do in fact confirm the idea that subfields of economics reach women and men in different ways. As an example, the survey data reveals that women are more likely than men to report that they find micro a more interesting subfield than macroeconomics, as well as less difficult and more intuitive.

In terms of the most natural occupations for economics graduates, our surveyed students view the economics profession as dominated by macroeconomics, and students entering economics with higher initial vocation also show a greater interest in macroeconomics. Further, the belief that the profession is male- and macro-dominated has a significant and appreciable impact on the probability of a woman declaring that she obtains worse grades in macro than in micro, while it does not influence the statements offered by men at all. As an additional piece of evidence that male students are viewed as the ones playing an influential role in macroeconomics, our peer effects analysis highlights that men are the only ones to influence their peers' performance in macroeconomics classes.

Taken together, the three pieces of evidence provided in this paper convey the idea that there is a gender imbalance across subfields in the profession that is not only evident at the research level but that starts as early as the undergraduate level. Different abilities do not seem to be the reason, while gender-based differences in interests and perceptions about the economics profession and the different subfields of our discipline might play an explanatory role.

In this paper we have not fully entered on the reasons why these preferences across subfields in economics differ by gender. These differences being natural or socially shaped remain a riddle to us. However, our observations suffice to conclude that the gender imbalance across subfields may be connected with the gender imbalance in economics as a whole: there are areas where women seem to feel relatively more comfortable, because of preferences or relative performance (or both), and these coincide with those that seem to be less well

known to our students. Our paper suggests that many microeconomics-related topics are not being effectively conveyed to young students - nor perhaps to the general public - and that correcting this misinformation might be one way to make economics a more appealing discipline for women.

Appendix A: JEL Approach

Table 6 : Correspondence between JEL codes and subfields

JEL code	Name	Area
A	General Economics and Teaching	Others
A0	General Economics and Teaching	Others
A1	General Economics	Others
A2	Economic Education and Teaching of Economics	Micro
A3	Collective Works	Others
B	History of Economic Thought, Methodology, and Heterodox Approaches	Others
B0	History of Economic Thought, Methodology, and Heterodox Approaches	Others
B1	History of Economic Thought through 1925	Others
B2	History of Economic Thought since 1925	Others
B3	History of Economic Thought: Individuals	Others
B4	Economic Methodology	Others
B5	Current Heterodox Approaches	Others
C	Mathematical and Quantitative Methods	Methods
C0	Mathematical and Quantitative Methods	Methods
C1	Econometric and Statistical Methods and Methodology: General	Methods
C2	Single Equation Models. Single Variables	Methods
C3	Multiple or Simultaneous Equation Models. Multiple Variables	Methods
C4	Econometric and Statistical Methods: Special Topics	Methods
C5	Econometric Modeling	Methods
C6	Mathematical Methods. Programming Models. Mathematical and Simulation Modeling	Methods
C7	Game Theory and Bargaining Theory	Micro
C8	Data Collection and Data Estimation Methodology. Computer Programs	Methods
C9	Design of Experiments	Micro
D	Microeconomics	Micro
D0	Microeconomics	Micro
D1	Household Behavior and Family Economics	Micro
D2	Production and Organizations	Micro
D3	Distribution	Micro
D4	Market Structure, Pricing, and Design	Micro
D5	General Equilibrium and Disequilibrium	Micro
D6	Welfare Economics	Micro
D7	Analysis of Collective Decision-Making	Micro
D8	Information, Knowledge, and Uncertainty	Micro
D9	Micro-Based Behavioral Economics	Micro
E	Macroeconomics and Monetary Economics	Macro
E0	Macroeconomics and Monetary Economics	Macro
E1	General Aggregative Models	Macro
E2	Consumption, Saving, Production, Investment, Labor Markets, and Informal Economy	Macro
E3	Prices, Business Fluctuations, and Cycles	Macro
E4	Money and Interest Rates	Macro
E5	Monetary Policy, Central Banking, and the Supply of Money and Credit	Macro
E6	Macroeconomic Policy, Macroeconomic Aspects of Public Finance, and General Outlook	Macro
E7	Macro-Based Behavioral Economics	Macro

JEL code	Name	Area
F	International Economics	Macro
F0	International Economics	Macro
F1	Trade	Macro
F2	International Factor Movements and International Business	Macro
F3	International Finance	Finance
F4	Macroeconomic Aspects of International Trade and Finance	Macro
F5	International Relations, National Security, and International Political Economy	Macro
F6	Economic Impacts of Globalization	Macro
G	Financial Economics	Finance
G0	Financial Economics	Finance
G1	General Financial Markets	Finance
G2	Financial Institutions and Services	Finance
G3	Corporate Finance and Governance	Finance
G4	Behavioral Finance	Micro
H	Public Economics	Micro
H0	Public Economics	Micro
H1	Structure and Scope of Government	Micro
H2	Taxation, Subsidies, and Revenue	Micro
H3	Fiscal Policies and Behavior of Economic Agents	Micro
H4	Publicly Provided Goods	Micro
H5	National Government Expenditures and Related Policies	Micro
H6	National Budget, Deficit, and Debt	Others
H7	State and Local Government. Intergovernmental Relations	Micro
H8	Miscellaneous Issues	Others
I	Health, Education, and Welfare	Micro
I0	Health, Education, and Welfare	Micro
I1	Health	Micro
I2	Education and Research Institutions	Micro
I3	Welfare, Well-Being, and Poverty	Micro
J	Labor and Demographic Economics	Micro
J0	Labor and Demographic Economics	Macro
J1	Demographic Economics	Others
J2	Demand and Supply of Labor	Micro
J3	Wages, Compensation, and Labor Costs	Micro
J4	Particular Labor Markets	Micro
J5	Labor-Management Relations, Trade Unions, and Collective Bargaining	Micro
J6	Mobility, Unemployment, Vacancies, and Immigrant Workers	Micro
J7	Labor Discrimination	Micro
J8	Labor Standards: National and International	Micro
K	Law and Economics	Micro
K0	Law and Economics	Others
K1	Basic Areas of Law	Micro
K2	Regulation and Business Law	Micro
K3	Other Substantive Areas of Law	Micro
K4	Legal Procedure, the Legal System, and Illegal Behavior	Micro

JEL code	Name	Area
L	Industrial Organization	Micro
L0	Industrial Organization	Micro
L1	Market Structure, Firm Strategy, and Market Performance	Micro
L2	Firm Objectives, Organization, and Behavior	Micro
L3	Nonprofit Organizations and Public Enterprise	Micro
L4	Antitrust Issues and Policies	Micro
L5	Regulation and Industrial Policy	Micro
L6	Industry Studies: Manufacturing	Micro
L7	Industry Studies: Primary Products and Construction	Micro
L8	Industry Studies: Services	Micro
L9	Industry Studies: Transportation and Utilities	Micro
M	Business Administration and Business Economics. Marketing. Accounting. Personnel Economics	Others
M0	Business Administration and Business Economics. Marketing. Accounting. Personnel Economics	Micro
M1	Business Administration	Others
M2	Business Economics	Others
M3	Marketing and Advertising	Others
M4	Accounting and Auditing	Others
M5	Personnel Economics	Others
N	Economic History	Others
N0	Economic History	Others
N1	Macroeconomics and Monetary Economics. Industrial Structure. Growth. Fluctuations	Macro
N2	Financial Markets and Institutions	Finance
N3	Labor and Consumers, Demography, Education, Health, Welfare, Income, Wealth, Religion, and Philanthropy	Micro
N4	Government, War, Law, International Relations, and Regulation	Others
N5	Agriculture, Natural Resources, Environment, and Extractive Industries	Others
N6	Manufacturing and Construction	Others
N7	Transport, Trade, Energy, Technology, and Other Services	Others
N8	Micro-Business History	Micro
N9	Regional and Urban History	Micro
O	Economic Development, Innovation, Technological Change, and Growth	Macro
O0	Economic Development, Innovation, Technological Change, and Growth	Macro
O1	Economic Development	Micro
O2	Development Planning and Policy	Micro
O3	Innovation. Research and Development. Technological Change. Intellectual Property Rights	Micro
O4	Economic Growth and Aggregate Productivity	Macro
O5	Economywide Country Studies	Macro
P	Economic Systems	Others
P0	Economic Systems	Macro
P1	Capitalist Systems	Others
P2	Socialist Systems and Transitional Economies	Others
P3	Socialist Institutions and Their Transitions	Others
P4	Other Economic Systems	Others
P5	Comparative Economic Systems	Others
Q	Agricultural and Natural Resource Economics. Environmental and Ecological Economics	Micro
Q0	Agricultural and Natural Resource Economics. Environmental and Ecological Economics	Micro
Q1	Agriculture	Others
Q2	Renewable Resources and Conservation	Micro
Q3	Nonrenewable Resources and Conservation	Micro
Q4	Energy	Micro
Q5	Environmental Economics	Micro
R	Urban, Rural, Regional, Real Estate, and Transportation Economics	Micro
R0	Urban, Rural, Regional, Real Estate, and Transportation Economics	Micro
R1	General Regional Economics	Others
R2	Household Analysis	Micro
R3	Real Estate Markets, Spatial Production Analysis, and Firm Location	Micro
R4	Transportation Economics	Micro
R5	Regional Government Analysis	Others

JEL code	Name	Area
Y	Miscellaneous Categories	Others
Y0	Miscellaneous Categories	Others
Y1	Data: Tables and Charts	Methods
Y2	Introductory Material	Others
Y3	Book Reviews (unclassified)	Others
Y4	Dissertations (unclassified)	Others
Y5	Further Reading (unclassified)	Others
Y6	Excerpts	Others
Y7	No Author General Discussions	Others
Y8	Related Disciplines	Others
Y9	Other	Others
Z	Other Special Topics	Others
Z0	Other Special Topics	Others
Z1	Cultural Economics. Economic Sociology. Economic Anthropology	Others
Z2	Sports Economics	Others
Z3	Tourism Economics	Others

Table 7 : Classification by themes of AEA papers

Cloud (main stems)	Rank	Share of Women 2014	2015	2016	Number of authors (average)
women, educ, gender, famili, children, age, gap, evid, parent, femal	1	48.40%	47.89%	46.75%	197
program, student, school, educ, learn, score, colleg, univers, effort	2	41.38%	32.26%	31.74%	155
health, insur, use, servic, provid, improv, examin, patient, impact, medic	3	31.63%	30.81%	34.13%	169
labor, employ, worker, wage, job, earn, skill, union, occup	4	28.25%	31.01%	29.65%	168
behavior, individu, experi, social, decis, choic, prefer, contribut, influenc	5	30.85%	24.24%	26.75%	170
econom, institut, will, social, theory, research, approach, understand, discuss	6	21.74%	28.49%	30.30%	169
effect, find, increas, result, impact, signific, suggest, decreas, variat, causal	7	36.11%	20.37%	22.78%	56
firm, competit, product, manag, ceo, busi, corpor, patent, innov, profit	8	25.00%	24.55%	27.39%	223
incom, tax, household, data, measur, wealth, consumpt, poverti, individu, level	9	29.82%	20.20%	19.90%	157
trade, countri, develop, import, product, export, global, foreign, china, region	10	20.39%	25.93%	21.77%	145

Classification by themes of AEA papers (continued)

Cloud (main stems)	Rank	Share of Women 2014	2015	2016	Number of authors (average)
financi, bank, credit, debt, market, risk, loan, mortgag, borrow	11	25.91%	21.55%	19.49%	308
growth, chang, declin, period, recess, year, rise, sever, trend, document	12	21.01%	30.86%	13.76%	103
state, govern, public, system, privat, legal, conflict, law, reform, polit	13	22.22%	20.00%	21.33%	86
capit, invest, cost, show, constraint, human, extern, rais, incent, lead	14	26.47%	12.50%	23.88%	69
use, local, chang, land, energi, citi, area, locat, environment, geograph	15	21.39%	15.17%	22.89%	205
polici, rate, shock, interest, monetary, exchang, real, chang, inflat, fiscal	16	23.64%	20.56%	15.24%	127
price, market, hous, demand, consum, use, valu, elast, real, properti	17	13.13%	20.63%	16.25%	160
inform, mechan, network, contract, agent, auction, result, strateg, signal, incent	18	18.69%	16.13%	14.47%	130
risk, return, fund, asset, stock, investor, volatil, hold, premium, equit	19	15.45%	16.17%	15.73%	249
estim, use, test, variabl, bias, identifi, distribut, error, result, data	20	12.96%	15.15%	15.50%	123
model, dynam, optim, welfar, heterogen, equilibrium, predict, general, endogen	21	7.34%	15.46%	11.00%	102

Appendix B: complementary material and tables.

Table 8 : Classification of courses at UV into subfields

	Core and Compulsory	Optional
Macroeconomics	Macroeconomics (principles) Macroeconomics (intermediate) Advanced Macroeconomics Dynamic Macroeconomics	International Macroeconomics. Topics in Applied Macroeconomics. Growth Theory.
Microeconomics	Microeconomics (principles) Microeconomics (intermediate) Advanced Microeconomics Games and Strategic Behaviour.	Industrial Organization. Applied Industrial Economics. Labour Economics. Topics in Applied Microeconomics. Game Theory. Natural Resources Theory. Microeconomics: Markets and Contracts.
Finance	Financial Accounting Financial Maths Principles of Financial Management Financial Information Analysis	Management in Banking and Finance. Business Finance. Theory of Banking and Finance. Financial and Monetary Policy. Financial Markets and Bank Operations.
Methods	Statistics Maths Mathematical Programming Econometrics (principles) Econometrics (intermediate)	Data Analysis in Economics. Introduction to Maths for Business. Econometric Applications. Maths for Dynamic Models.
Other	Basics of Law European Union Economy Economic History Foundations of Business Management Accounting Economic Policy (principles) Economic Policy (advanced) International Economics Public Economics	Regional Economics. International Trade. Competitiveness and R&D Policies. Growth and Development in Ec. History. Agricultural Economics. Social Economics. Public Economics Evaluation. Economic Policy in the UE. Public Regulation. Tax System in Spain. Regional and Local Development Policies. Monetary Integration. Business Strategies. Federal Taxation. Contemporary Economic Thinking .

Courses difficult to categorize by their subject matter have been classified according to the fields of the departments that are in charge of the course.

Appendix C: Two-steps Semi-parametric selection model.

To account for selection bias into optional courses we start defining a dummy variable s_i which takes the value 1 if the student has chosen any optional course in the f field for which we estimate the equation. The QR equation for grades can be written as follows:

$$Q_q(G_{imt}/s_i = 1) = \alpha(q) + \beta(q)W_i + \gamma(q)\mathbf{x}_{it} + m_t(q) + \tau(q) + Q_q(u1_{imt}/s_i = 1) \tag{4}$$

where now the quantiles are conditional on selection. The selection equation follows a semi-parametric single index regression model as:

$$s_i = g(X_i'\beta_s) + u2_i \tag{5}$$

where the function g is of unknown form and $(X_i'\beta_s)$ is a single linear combination of parameters and covariates thought to explain the selection. A student may choose none, one or more than one course in a given subfield. Students in two courses of a given subfield will contribute twice to the QR regression of this subfield, and will also appear twice in the selection equation in that subfield. This yields an estimation which is somewhat equivalent to applying frequency weights, that is, weighting individuals' observations according to their frequency in the sample.

The conditional quantile of observed grades depends on the quantile-specific selection bias term of unknown form $g(X_i'\beta_s)$ which cannot be corrected using the traditional parametric estimation *à la* Heckman.[36] Instead, we use the two-step procedure suggested by Buchinsky (1998) as follows. In the first step, the parameters in the selection equation (5) are estimated using any semi non-parametric estimator. As in Sharma et al. (2013), we use the SNP estimator proposed by Gallant and Nychka (1987), which has the advantage over other semi-parametric estimators that it does not involve kernel smoothing.[37] In the second step, the QR is run on all the second stage covariates and on a polynomial expansion on the estimated linear index $(X_{if}'\beta_s)$ of the first stage. The degree of the polynomial in our QR estimation has been chosen according to the statistical significance obtained for the different

[36]Parametric estimators are consistent and asymptotically efficient only if the distributional assumptions are correctly specified. Since in QR the selection bias term is of unknown form, its parametric estimation would lead to inconsistent estimates.

[37]Kernel-based Methods require the choice of bandwidth, are computationally intensive and often have convergence difficulty (Stewart, 2004). The SNP is a pseudo-ML estimator where the unknown likelihood is approximated using a Hermite polynomial expansion. Identification of an SNP estimator requires fixing the value of the intercept and having at least one continuous variable in the vector of covariates. Following Melenberg and Soest (1996) the value of intercept is set to its probit estimate. Also, continuous variables are included in the estimation. See De Luca (2008) for details on the command SNP in Stata.

powers of the index.

The linear index estimated in our SNP model comprises all the covariates already included in the QR tables, including the women dummy variable. We also add as determinants of the choice of optional courses the average grades students obtained previously in the different compulsory courses. This reflects the plausible idea that students may choose optional courses on the basis of how good they are at the different subfields.[38] All covariates that refer to the students themselves (age, full-time study, high school type, etc.) as well as the grades previously obtained in the compulsory courses are interacted with the women dummy to allow us to observe whether women make their choices differently to men.

Parents' educational and employment status are not crossed with the women dummy, after having checked that no differences between men and women arose in any case in the estimated effect of these variables. These variables are not presented in the table for the sake of simplicity (and, in general, do not show significant effects in most of the cases).[39]

We apply the two-step procedure to each of the subfields. That is, for macro, micro, finance, Methods and other, we estimate the non-linear selection equation and, then, the corresponding QR regression with the selection correction terms. Table 9 displays the results for the SNP estimation, while Table 10 presents the result for the QR of grades. In Table 9 a double column has been constructed that shows the baseline effect of each covariate, plus the differential effect for women. That is, the second column, labeled women (difference), in each subfield reports the coefficient of the interaction term although there is only one estimated equation in each pair of columns.

One of the most interesting findings from the SNP estimation is the different impact of past grades in macro and micro on the students' choice of the corresponding optional courses. On the one hand, the better the past grades in macro, the higher the likelihood of a student choosing courses in the macro, micro and finance subfields. No differences arise between men and women in this regard. However, past performance in micro seems to affect the choice of micro courses differently in men and women. In particular, women with good marks in micro are far less likely to choose macro courses than men: the negative coefficient for the interaction with the women dummy in the second column for macro (-1.196) almost cancels out the positive baseline effect (1.582). In microeconomics, good grades in the past make women much more prone than men to choose micro-related optional courses: the estimated

[38]Performance in previous economic courses by high school students is found to be a positive predictor of persistence in the discipline (Chizmar, 2000; Jensen and Owen, 2001; Calkins and Welki, 2006).

[39]It could be argued that the decisions to enroll in different fields may be interconnected. Then, the selection model could be thought of as a simultaneous equation system where the choice of one subfield depends on the previous results in that field and on the choice made about the rest of the subfields, plus additional covariates. By running separate equations, each of them can be considered as the reduced form equation from such a joint system, where the past grades in all courses enter into each of the equations. The efficiency gain from the joint estimation of the reduced form equations is not a primary goal since our first step SNP is basically aimed at estimating the index to correct for selection in the second step.

effect for men is 1.253 while for women an additional 0.950 is significantly estimated.

The selection into finance also shows some differences between men and women. In particular, it seems that men are not driven by their past marks in finance to enroll in finance-related courses, whereas women are more likely to enroll in finance courses if they previously achieved good grades in finance. Another notable result is the negative association between good marks in macro and micro and the likelihood of enrolling in the subfields of Methods or other. Women with previous good grades in micro are much less likely than men to enroll in Methods, while they pay much more attention than men to their past marks in Methods when selecting courses in this subfield.

Additional results in Table 9 refer to the association between the AU-score and the choice of courses. Higher ability students, as proxied by the AU-score, tend to choose macro and Methods to a greater extent, while they seem to be less attracted by finance. In the rest of the subfields, no significant effects are estimated, either for men or for women.

The women dummy variable coefficient is positive in macro and Methods, which accounts for the part of the probability that remains unexplained for women as compared to men after partialing out the impact of the rest of covariates. The estimated total probability, on average, that women enter a given subfield as compared to men is not, however, uniquely defined by this dummy coefficient. To offer some results in this regard, at the bottom of Table 9 we show the average value of the estimated index ($X_i'\beta_s$) for men and the differential value for women (in particular, we run an OLS regression of the index on a constant and the women dummy). The obtained estimates indicate that the index is significantly lower for women in macro, finance and Methods while the difference is positive in micro and other. As the probability increases with the index, these results indicate that women have lower estimated probabilities of entering macro, finance and Methods, as compared to micro and other courses.

Table 9 : Semi-Non-Parametric (SNP) estimation of optional courses choices

	Macro		Micro		Finance		Methods		Other	
	men	women (difference)	men	women (difference)	men	women (difference)	men	women (difference)	men	women (difference)
Women		1.636***		0.282		0.194		2.346***		-0.220
		(0.533)		(0.793)		(0.277)		(0.376)		(0.356)
AU-score	0.0654**	-0.0494	0.0742	-0.0229	-0.0893***	0.0745*	0.185***	-0.208***	0.00983	0.0123
	(0.0295)	(0.0411)	(0.110)	(0.116)	(0.0239)	(0.0389)	(0.0381)	(0.0530)	(0.0241)	(0.0665)
Past grades macro	2.570***	0.241	2.627***	0.0410	2.839***	0.0461	-1.689***	-0.203	-1.536***	-0.0496
	(0.158)	(0.278)	(0.419)	(0.540)	(0.195)	(0.313)	(0.527)	(0.724)	(0.306)	(0.497)
Past grades micro	1.582***	-1.196***	1.253***	0.950*	1.665***	0.0998	-2.893***	-2.595***	0.0929	0.0413
	(0.162)	(0.397)	(0.382)	(0.501)	(0.144)	(0.356)	(0.451)	(0.606)	(0.256)	(0.381)
Past grades finance	-0.224	0.488	2.515***	-2.055***	0.100	0.736*	-0.770	0.877	-0.175	0.893*
	(0.369)	(0.489)	(0.434)	(0.554)	(0.228)	(0.406)	(0.513)	(0.623)	(0.396)	(0.491)
Past grades Methods	0.0639	-0.824	1.501***	-0.748	-0.371	-0.156	1.500***	1.168***	-1.276***	-0.432
	(0.370)	(0.519)	(0.285)	(0.805)	(0.228)	(0.313)	(0.0876)	(0.283)	(0.252)	(0.316)
Past grades other	0.00936	0.368*	0.0720	0.374	0.229**	-0.0425	0.367***	-0.583***	-0.172*	0.0262
	(0.127)	(0.196)	(0.268)	(0.393)	(0.108)	(0.156)	(0.140)	(0.198)	(0.101)	(0.149)
Age	0.0213***	-0.0766***	-0.0344***	-0.0421*	0.0101**	-0.0236***	-0.0190***	-0.0410**	0.0126	0.0189
	(0.00531)	(0.0203)	(0.0129)	(0.0241)	(0.00407)	(0.00907)	(0.00613)	(0.0165)	(0.0226)	(0.0123)
Living at home	-0.0544	0.0295	-0.399	0.239	-0.00370	0.0260	0.0982	-0.0259	0.0767	-0.0823
	(0.0938)	(0.128)	(0.367)	(0.448)	(0.0644)	(0.0950)	(0.0920)	(0.126)	(0.0600)	(0.0790)
Standard high school	0.175	0.164	-0.135	0.320	-0.377***	0.175	1.192***	-1.195***	0.104	-0.112
	(0.131)	(0.175)	(0.425)	(0.552)	(0.0946)	(0.152)	(0.194)	(0.245)	(0.0917)	(0.151)
Vocational high school	0.0818	0.569**	-0.807	0.489	-0.0445	0.193	0.0552	-0.476	0.0202	-0.295*
	(0.178)	(0.235)	(0.516)	(0.639)	(0.130)	(0.187)	(0.248)	(0.343)	(0.140)	(0.168)
full-time student	0.166*	-0.118	0.146	0.157	0.0717	-0.150	-0.0337	0.0282	0.0244	0.0185
	(0.0966)	(0.154)	(0.359)	(0.409)	(0.0681)	(0.106)	(0.0933)	(0.148)	(0.0637)	(0.111)
Constant (fixed)	-3,218		-1,311		-2,099		-1,560		1,138	
Likelihood Ratio Test for probit vs. SNP [p-values in brackets]	37.93 [0.000]		4.61 [0.032]		106.31 [0.000]		186.82 [0.000]		13.24 [0.000]	
Log pseudolikelihood	-1199.92		-5076.24		-2716.64		-3579.03		-8101.24	
	men	women (difference)	men	women (difference)	men	women (difference)	men	women (difference)	men	women (difference)
Estimated linear index	1.154***	-0.021***	0.002	0.018*	0.394***	-0.122***	0.843***	-0.188***	0.012***	0.046***
	(0.004)	(0.006)	(0.008)	(0.010)	(0.,005)	(0.007)	(0.007)	(0.011)	(0.003)	(0.005)

Robust errors in parentheses. * p<0.10 ** p<0.05 *** p<0.01.

Table 10 : QR of grades in optional courses by subfield (with and without selection correction)

	Macro No corr.	Macro Select corr.	Micro No corr.	Micro Select corr.	Finance No corr.	Finance Select corr.	Methods No corr.	Methods Select corr.	Other No corr.	Other Select corr.
Women	-0.521	-0.584*	0.364***	0.352***	0.102	0.180	0.078	0.014	-0.028	-0.019
	(0.333)	(0.332)	(0.087)	(0.106)	(0.166)	(0.168)	(0.138)	(0.143)	(0.032)	(0.031)
AU-score	0.049	-0.091	0.226***	0.092**	0.177***	0.184***	0.208***	0.279***	0.119***	0.094***
	(0.136)	(0.139)	(0.048)	(0.040)	(0.056)	(0.052)	(0.054)	(0.056)	(0.013)	(0.014)
W*AU-score	0.129	0.155	-0.098**	-0.074	-0.007	-0.028	-0.030	-0.067	-0.004	-0.001
	(0.189)	(0.175)	(0.041)	(0.050)	(0.083)	(0.072)	(0.067)	(0.068)	(0.019)	(0.017)
Age	-0.042**	-0.057***	-0.014**	0.004	-0.012	-0.011	0.001	-0.008	-0.009***	-0.002
	(0.016)	(0.017)	(0.007)	(0.011)	(0.013)	(0.012)	(0.014)	(0.014)	(0.003)	(0.003)
Standard high school	0.322	-0.015	0.326	0.074	0.175	0.335*	0.132	0.623**	0.233***	0.175***
	(0.693)	(0.702)	(0.198)	(0.155)	(0.183)	(0.192)	(0.263)	(0.246)	(0.051)	(0.055)
Vocational high school	-0.396	-0.521	0.298	0.416**	0.325	0.284	0.164	0.185	0.165***	0.106*
	(0.772)	(0.720)	(0.233)	(0.198)	(0.223)	(0.224)	(0.289)	(0.279)	(0.063)	(0.059)
full-time student	0.432	0.208	0.289*	0.180	0.017	0.015	0.246*	0.216	0.125***	0.127***
	(0.273)	(0.282)	(0.168)	(0.116)	(0.172)	(0.169)	(0.136)	(0.144)	(0.030)	(0.031)
Living at home	-0.117	-0.052	0.018	0.151	0.127	0.124	-0.006	0.005	-0.014	0.006
	(0.243)	(0.213)	(0.064)	(0.097)	(0.134)	(0.129)	(0.115)	(0.113)	(0.024)	(0.024)
SNP index		1.330***		0.645***		0.946***		-0.567***		-0.550***
		(0.280)		(0.087)		(0.256)		(0.134)		(0.050)
SNP index ^{2}				0.293***				0.213***		0.456***
				(0.065)				(0.048)		(0.066)
SNP index ^{3}				-0.096***						0.215***
				(0.025)						(0.042)
Constant	9.477***	8.626***	5.702***	5.579***	6.422***	5.694***	5.699***	5.710***	5.061***	4.994***
	(1.363)	(1.359)	(0.630)	(0.408)	(0.799)	(0.762)	(0.565)	(0.577)	(0.176)	(0.160)
N Obs	256	256	1624	1624	755	755	1062	1062	12734	12734
Adj. R^2	0.176	0.256	0.051	0.115	0.071	0.128	0.069	0.091	0.426	0.437

Bootstrapped errors in parentheses (500 replications). * p<0.10 ** p<0.05 *** p<0.01. All regressions include course-year effects, year dummies and the complete set of socio-economic controls included in Table 1. AU-score: student's score to access university (averaged grades from high school and the university access exam).

Appendix D: IVs estimation of contagion or peer effects.

To overcome the problem of endogeneity of the peer effects measures in equations 4 and 5 we follow the most widely-used approach, which consists in using instrumental variables (e.g., Goux et al., 2007). These instruments may be constructed from individual measures affecting the individual responses about interests and grades which are exogenous to the error term of the equation. We choose two variables that are predetermined before students enter their respective groups and before their interaction in the classroom starts: the AU-scores and their self-reported statements about initial vocation. With the individual responses to these two variables, we construct the IVs.

More specifically, we define for each student the percentages of males and female peers in the class expressing an initial vocation PV_{j-i} and their respective average AU-scores PAU_{j-i}. In a first stage, we estimate the reduced forms of the male peers' interest and the female peers' interest on the IVs and the rest of the covariates of the main equation. As an example, for the female peer measure in equation 2 the first stage regression takes the general form:

$$PI_i^W = \beta_0 + \beta_1 PAU_{j-i}^W + \beta_2 PAU_{j-i}^M + \beta_3 PV_{j-i}^W + \beta_4 PV_{j-i}^M + \beta_5 \mathbf{w}_i + e_i^I \tag{6}$$

where $\mathbf{w}_{is}$ stands for all the covariates and controls other than the peer measures appearing in the main equation (equation 2 in this case). An equation like 6 is run for each of the four peer measures appearing in equations 2 and 3 above. In the second stage, we plug the first-stage predictions for the peer measures into the right-hand side of the corresponding main equation of interest. To account for the extra variance introduced in the second stage when a first-stage estimated variable is included, we bootstrap the standard errors.

The application of IV to estimate peer effects present a fundamental difference with respect to the IV estimation in the standard setting. Here, the individual variable used to construct the IV does not fulfill an exclusion restriction in the main equation. On the contrary, if an individual covariate is not explanatory of the individual responses its aggregation will probably not correlate with the peer measure. Moreover, the individual variable may enter the main equation together with their peer construction. An example of an application of this approach can be found in Carrell et al. (2008) for the estimation of peer effects in academic cheating.

References

Bagues, M., Campa, P. (2018). "Can Gender Quotas in Candidate Lists Empower Women? Evidence from a Regression Discontinuity Design". *CEPR Discussion Paper* No. 12149,

Bayer, A., and Rouse, C. E. (2016). "Diversity in the economics profession: A new attack on an old problem". *Journal of Economic Perspectives*, 30(4), 221-42.

Blei, D. M., Ng, A.Y., Jordan, M. I. (2003). "Latent Dirichlet Allocation" *The Journal of Machine Learning Research*, 3, 993-1022.

Boring, A., Ottoboni, K., and Stark, P. (2016). "Student evaluations of teaching (mostly) do not measure teaching effectiveness". *ScienceOpen Research*.

Buchinsky, M. (1998). "The dynamics of changes in the female wage distribution in the USA: a quantile regression approach". *Journal of Applied Econometrics*, 13(1), 1-30.

Calkins, L.N. and A. Welki. 2006. "Factors that influence choice of major: Why some students never consider economics". *International Journal of Social Economics*, 33(8): 547-564.

Chari, A. and Goldsmith-Pinkham, P. (2017). "Gender Representation in Economics Across Topics and Time: Evidence from the NBER Summer Institute", NBER Working Papers 23953.

Chizmar, J. F. (2000). "A discrete-time hazard analysis of the role of gender in persistence in the economics major". *Journal of Economic Education*, 31(2), 107-118.

Della Giusta, M., Jewell, S., and Greetham, D. V. (2017). "Beliefs, Exams and Social Media: A Study of Girls and Boys in the UK". Economics and Management Discussion Papers, em-dp2017-02. Henley Business School, Reading University.

De Luca, G. (2008). "SNP and SML estimation of univariate and bivariate binary-choice models". *Stata Journal*, 8(2), 190-220.

Dolado, J. J., Felgueroso, F., and Almunia, M. (2012). "Are men and women-economists evenly distributed across research fields? Some new empirical evidence". *SERIEs*, 3(3), 367-393.

Emerson, T. L., McGoldrick, K., and Mumford, K. J. (2012). "Women and the choice to study economics". *The Journal of Economic Education*, 43(4), 349-362.

Ferber, M. A. (1995). "The study of economics: A feminist critique". *The American Economic Review*, 85(2), 357-361.

Gallant, A.R., D.W. Nychka (1987). "Semi-nonparametric maximum likelihood estimation". *Econometrica*, 55(2), 363-390.

Glaeser, E. L., Sacerdote, B. I., and Scheinkman, J. A. (2003). "The social multiplier". *Journal of the European Economic Association*, 1(2-3), 345-353.

Gneezy, U., and Rustichini, A. (2004). "Gender and competition at a young age". *American Economic Review*, 94(2), 377-381.

Goux, D., and Maurin, E. (2007, October). "Close neighbours matter: Neighbourhood effects on early performance at school". *Economic Journal*, 117(523), 1193-1215.

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998). "Measuring individual differences in implicit cognition: the implicit association test". *Journal of Personality and Social Psychology*, 74(6), 1464-80.

Grunspan, D. Z., Eddy, S. L., Brownell, S. E., Wiggins, B. L., Crowe, A. J., and Goodreau, S. M. (2016). "Males under-estimate academic performance of their female peers in undergraduate biology classrooms". *PloS one*, 11(2): e0148405

Hansen, S., McMahon, M., Prat, A. (2018). "Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach". *The Quarterly Journal of Economics*, 133(2), 801-870.

Heckman, J.J. (1979). "Sample selection bias as a specification error". *Econometrica*, 47(1), 153-161.

Hengel, E. (2018) "Publishing while female. Are women held to higher standards? Evidence from peer review". Third chapter of the author's dissertation, University of Cambridge, September 2015.(http://www.erinhengel.com/research/publishing-female.pdf, last accessed 09 Nov. 2018.)

Hoogendoorn, S., Oosterbeek, H., and Van Praag, M. (2013). "The impact of gender diversity on the performance of business teams: Evidence from a field experiment". *Management Science*, 59(7), 1514-1528.

Huntington-Klein, N., and Rose, E (2018), "Gender Peer Effects in a Predominantly Male Environment: Evidence from West Point". *American Economic Review*, Papers and Proceedings, 108, 392-395.

Iriberri, N., and Rey-Biel, P. (2017). "Stereotypes are only a threat when beliefs are reinforced: On the sensitivity of gender differences in performance under competition to information provision". *Journal of Economic Behavior and Organization*, 135, 99-111.

Jensen, E. J., and Owen, A. L. (2001). "Pedagogy, gender, and interest in economics". *Journal of Economic Education*, 32(4), 323-343.

Lundeberg, M. A., Fox, P. W., and Punćcohaŕ, J. (1994). "Highly confident but wrong: Gender differences and similarities in confidence judgments". *Journal of Educational Psychology*, 86(1), 114.

MacNell, L., Driscoll, A., and Hunt, A. N. (2015). "What's in a name: Exposing gender bias in student ratings of teaching". *Innovative Higher Education*, 40(4), 291-303.

McPherson, M., Smith-Lovin, L. and Cook, J. M. (2001) "Birds of a Feather: Homophily in Social Networks", *Annual Review of Sociology* 27(1): 415-444.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman,

J. (2012). "Science faculty's subtle gender biases favor male students". *Proceedings of the National Academy of Sciences*, 109(41), 16474-16479.

Mueller, H., Rauh, C, (2017). "Reading Between the Lines: Prediction of Political Violence Using Newspaper Text". *American Political Science Review*, 112(2), 358-375.

Phelps, E. (1972). "The Statistical Theory of Racism and Sexism", *American Economic Review* 62(4), 659-661.

Rask, K. and Tiefenthaler, J. 2008. "The role of grade sensitivity in explaining the gender imbalance in undergraduate economics". *Economics of Education Review*, 27, 676-687.

Sarsons, H. (2017). "Recognition for group work: gender differences in academia". *American Economic Review Papers & Proceedings*, 107(5), 141-145.

Sharma, A., Siciliani, L., and Harris, A. (2013). "Waiting times and socioeconomic status: does sample selection matter?". *Economic Modelling*, 33, 659-667.

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). "Stereotype threat and women's math performance". *Journal of Experimental Social Psychology*, 35(1), 4-28.

Steele, C. M., and Aronson, J. (1995). "Stereotype threat and the intellectual test performance of African Americans". *Journal of Personality and Social Psychology*, 69(5), 797.

Stewart, M.B. (2004). "Semi-nonparametric estimation of extended ordered probit models". *Stata Journal*, 4(1), 27-39.

Tang, C, Ross K., Saxena, N. and Chen, R. (2011). "What's in a name: a study of names, gender inference, and gender behavior in facebook". *Database Systems for Advanced Applications*, 344-356.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Malone, T. W. (2010). "Evidence for a collective intelligence factor in the performance of human groups". *Science*, 330(6004), 686-688.

Wu, A. H. (2018). "Gendered Language on the Economics Job Market Rumors Forum". *American Economic Review: Papers & Proceedings* 108, 175-79.