

Rethinking evolutionary economic geography through a gender lens

Barbara Martini* 

Business Engineering, University of Rome Tor Vergata, 1 00133, Rome, Italy

*Corresponding author. Business Engineering, University of Rome Tor Vergata, Viale del Politecnico, 1 00133, Rome, Italy.
E-mail: barbara.martini@uniroma2.it

Abstract

Evolutionary Economic Geography (EEG) explores how industries cluster, networks evolve, and regions grow, yet often overlooks gender. This study introduces a gender-sensitive framework, extending Related and Unrelated Variety (RV and UV) measures through a three-stage entropy decomposition. Findings reveal that RV and UV respond differently as women's employment share increases, showing increasing returns when women are underrepresented. Thus, promoting women's participation in male-dominated industries can more effectively enhance labor market dynamics than concentrating them in already female-dominated sectors. This underscores the importance of not only encouraging women's labor market participation but also considering the sectors in which they are employed.

Keywords: gender and knowledge spillovers; related variety; unrelated variety; regional economic growth; labor market composition; evolutionary economic geography..

JEL classifications: R11; O33; J16

Highlights

- Gender differences affect the dynamics of knowledge spillovers across industries.
- Traditional measures of Related and Unrelated Variety overlook gender dimensions.
- A three-stage decomposition is proposed to capture gender-based knowledge spillovers within and between industries.
- Theoretical results show that the gender composition of industries significantly shapes the behavior of knowledge spillovers.
- Increasing female employment in male-dominated industries is more effective for enhancing knowledge diffusion and regional growth.

1. Introduction

According to the definition provided by the World Health Organization¹, gender refers to the socially constructed characteristics and roles of women and men, while sex refers to biologically determined traits. Individuals are born female or male but learn to become girls and boys who grow into women

¹ <https://www.who.int/europe/health-topics/gender>

and men. This learned behavior shapes gender identity and defined gender roles. Economic theory began incorporating gender through early contributions such as [Hakim \(2000\)](#), who introduced preference theory to explain women's choices between paid employment and family work.

Despite a growing body of empirical literature on gender in labor economics, Evolutionary Economic Geography (EEG) has yet to meaningfully integrate gender into its conceptual and empirical frameworks. In labor economics, research has extensively documented the persistence of gender gaps in wages and employment outcomes ([Blau and Kahn 2017](#)) as well as the broader economic costs of talent misallocation due to gendered barriers in labor markets ([Hsieh et al., 2019](#)). However, these insights have rarely been connected to EEG's core concerns—namely, knowledge spillovers and spatial development. Recent work in economic geography has emphasized the role of institutions and inclusivity in shaping development outcomes. [Rodríguez-Pose and Ketterer \(2020\)](#) and [Rodríguez-Pose and Ganau \(2022\)](#) show that institutional exclusion fosters long-term underperformance and discontent. [Giannantoni and Rodríguez-Pose \(2025\)](#) further demonstrate that the institutional capacity of regional governments significantly influences women's ability to reconcile career and motherhood, affecting both labor participation and regional growth. More broadly, [Rodríguez-Pose \(2018\)](#) has argued that regions failing to include their populations in development processes risk becoming the so-called “places that don't matter.” At the same time, research on economic complexity highlights that innovation increasingly depends on the recombination of knowledge in diverse environments. [Balland et al. \(2020\)](#) show that complex activities concentrate in large cities not simply because of their size, but because of their potential for knowledge recombination. Excluding women from these recombination spaces means leaving untapped a significant share of local innovation capacity. A growing body of economic research has investigated differences in the behavior of women and men in the labor market and the resulting consequences for gender segregation. Gender segregation—defined as the dominance of one sex in a particular occupation, or the overrepresentation of one sex relative to its expected share—can be horizontal or vertical. Horizontal segregation typically refers to the unequal distribution of women and men across industries and occupations, while vertical segregation reflects disparities in career progression, status, and leadership access. Segregation often constrains women's participation in the labor market. However, higher participation rates do not necessarily reflect greater gender equality or female empowerment. Due to the uneven distribution of women and men across sectors, increases in female labor force participation tend to be concentrated in industries where women are already overrepresented. [Ngai and Petrongolo \(2017\)](#) argue that women have a comparative advantage in service production. In Italy, the evolution of the production structure—particularly the process of de-specialization since 1995 ([Martini 2020](#))—has increased employment in services, thus boosting women's employment in these sectors ([Olivetti and Petrongolo 2016](#); [Petrongolo and Ronchi 2020](#)). Addressing occupational segregation is essential to closing the gender wage gap, improving job quality and earnings, and increasing women's labor force participation. Beyond the individual level, employment segregation also affects economic growth, household welfare, firm performance, and intergenerational social mobility. Reducing segregation can generate a virtuous cycle, in which increased female participation in high-return occupations fosters stronger professional networks and challenges prevailing social norms ([Das and Kotikula 2018](#)). Nonetheless, women's labor market outcomes continue to be shaped by institutional and cultural factors, including welfare regimes, employment protection policies ([Hall and Krueger 2018](#)), and social norms ([Alesina et al., 2013](#)).

Despite rising interest in integrating gender into various branches of economic analysis, gender remains largely invisible in EEG models—even though it significantly shapes how knowledge flows. How can gender—through differences in employment segregation, social networks, and innovation behavior—reshape the way we understand proximity and knowledge spillovers in Evolutionary Economic Geography? Traditional EEG models assume gender homogeneity. But how limiting is that assumption in explaining real-world patterns? If gender shapes knowledge diffusion, can we continue to rely on gender-blind frameworks in EEG? Furthermore, EEG is concerned with the uneven spatial distribution of economic activities, focusing on the historical processes that produce these patterns, with [Nelson and Winter's \(1982\)](#) concept of organizational routines as the central unit of analysis ([Boschma and Frenken 2011](#)). As [Ter Wal and Boschma \(2011\)](#) argue, firms differ widely in their capabilities, strategies, and routines.

Yet Nelson and Winter's routines assume homogeneous human capital. In practice, individuals within firms are heterogeneous: they differ in values, beliefs, preferences—and gender ([Felin et al., 2012](#)). These differences are central to understanding organizational performance and innovation ([Abel et al., 2008](#); [Grant 1996](#); [Simon 1991](#)). Empirical evidence shows that gender diversity enhances

knowledge outcomes (Nielsen et al., 2017; Valantine and Collins 2015) and fosters innovation by introducing diverse perspectives (Sastre 2015; Østergaard et al., 2011; Jackson et al., 1995; O'Reilly et al., 1997; Xie et al., 2020). Gender diversity also brings social benefits by contributing varied value systems and behavioral tendencies—for example, men tending to be more assertive and task-oriented, while women are often more process-focused and collaborative (Karakowsky and Siegel 1999; Wood et al., 1985). Meta-analyses confirm that gender-diverse groups outperform homogeneous ones (Bowers et al., 2000; Williams and O'Reilly 1998; Wood et al., 1985). More recent research shows that gender-diverse boards are linked to higher patenting rates, more novel innovations, and greater innovation efficiency (Griffin et al., 2021), while women in management roles significantly contribute to R&D and organizational innovation (Hernández-Lara et al., 2021; Torchia et al., 2018). Consequently, the assumption of homogeneity in Nelson–Winter routines should be revised to reflect gender heterogeneity. In the context of industry relatedness and knowledge spillovers, a central question arises: do firms learn more from local firms in the same industry (regional specialization) or from local firms in other industries (regional diversity)? In other words, are the most innovative and fast-growing regions sector-specialized or diversified? Although regional diversification is widely seen as critical for creating new growth paths and offsetting stagnation (Boschma and Gianelle 2014), the causal relationship between regional industrial structure and growth remains ambiguous. While Marshall–Arrow–Romer (MAR) externalities emerge from intra-industry interactions, Jacob's externalities derive from inter-industry spillovers. From this perspective, knowledge exchange across sectors is crucial for innovation, and economic diversity is a key driver of urban prosperity. However, as Jacobs (1961, 1969) emphasized, diversity alone is not sufficient—effective complementarities and shared “pools of use” are required. When cognitive distance is too great, recombination becomes ineffective (Nooteboom 2000; Aarstad et al., 2016). By contrast, knowledge flows between related sectors are more likely to enable novel recombination and stimulate innovation. The literature has conceptualized industry relatedness through the notions of related and unrelated variety, based on the work of Boschma (2005), Frenken et al. (2007), and Boschma and Iammarino (2009), using Theil's entropy index (1965, 1967). Related variety enables Jacobs-type externalities across cognitively similar sectors, promoting innovation. Unrelated variety, meanwhile, captures cross-sectoral diversity that—according to portfolio theory—enhances regional resilience and mitigates labor market vulnerability to shocks. Frenken et al. (2007) developed a two-level entropy-based decomposition to differentiate between these two forms of variety.

In conclusion, integrating gender into EEG requires rethinking its micro-foundations by relaxing the assumption of individual homogeneity. Proximities are shaped not only by industrial and spatial factors but also by gender segregation: men and women are unevenly distributed across industries, which influences both relatedness and knowledge diffusion. As such, industry relatedness itself may need to be reconceptualized. Innovation and regional growth trajectories may be accelerated—or constrained—depending on the degree of gender equality in labor force participation and the extent of occupational segregation. In this contribution, we extend Frenken et al. (2007) entropy-based decomposition to incorporate a gender dimension. Specifically, we propose a three-stage decomposition that yields five indicators: one capturing overall gender diversity in the economy, and four measuring related and unrelated variety by gender. This refined approach not only improves measurement precision but also exposes structural asymmetries in how knowledge flows within gendered labor markets. To our knowledge, this is the first study to examine the intersection of gender, knowledge spillovers, and regional growth using entropy-based indicators. The remainder of this paper is structured as follows. Section 2 outlines the conceptual framework of knowledge spillovers and industry relatedness. Section 3 presents the standard entropy-based decomposition of related and unrelated variety. Section 4 introduces our gender-sensitive three-stage decomposition and describes the construction of five novel entropy indicators. Section 5 reports the main empirical findings. Finally, Section 6 discusses policy implications and offers directions for future research.

2. From variety to related and unrelated variety: one stage decomposition

Related and Unrelated Variety (Boschma 2005; Frenken et al. 2007; Boschma and Iammarino 2009) are derived from Theil's index (1967), an inequality measure rooted in Shannon's (1948) entropy. The

Table 1. Glossary of symbols and measures.

Symbol	Meaning	Definition
j	Region	Index for territorial unit.
$K \in \{F, M\}$	Gender	Female (F) or Male (M).
G	g-digit industry	Intermediate industry level.
i	i-digit industry	Detailed industry level within g.
G	Numbers of g-digit industries	$G = \sum g.$
N	Numbers of i-digit industries	$N = \sum i.$
E_{ij}	Employment in i-digit i in region j	Level within g to which i belongs.
E_{gj}	Employment in g-digit g in region j	Sum over $i \in g$ of E_{ij}
E_{Kj}	Employment of gender K	$E_j = E_{Fj} + E_{Mj}$
E_{Kg}	Employment of gender K in 2-digit g	Sum over $i \in g$ of E_{Kij}
E_j	Total employment in region j	Sum over all i (or g, K).
p_{ij}	Share of i in region j	$p_{ij} = E_{ij}/E_j$
p_{Kij}	Share of K and i in region j	$p_{Kij} = E_{Kij}/E_j$
p_{gj}	Share of g in region j	$p_{gj} = E_{gj}/E_j$
p_{Kgj}	Share of K and g in region j	$p_{Kgj} = E_{Kgj}/E_j$
P_{Kj}	Gender share in region j	$P_{Kj} = E_{Kj}/E_j$
VAR	Total variety over i-digit	$VAR = -\sum_{i=1}^N P_i \ln(P_i)$
UV	Unrelated variety across g-digit	$UV = [\sum_{g=1}^G P_g \ln \frac{1}{P_g}]$
RV	Related variety within g (over $i \in g$)	$RV = \sum_{g=1}^G P_g H_g$
UV_k	Diversity between genders	$UV_k = \sum_{k=1}^2 P_k \ln(\frac{1}{P_k})$
UV_{kg}	Diversity of g within each gender	$\sum_{k=1}^2 P_k H_k$ where $H_k = \sum_{g=1}^G \frac{P_{kg}}{P_k} \ln(\frac{P_k}{P_{kg}})$
RV_K	Within-(g, K) diversity over i	$\sum_{g=1}^G P_{kg} H_{kg}$ where $H_g = [\sum_{i \in g} \frac{P_i}{P_g} \ln(\frac{P_g}{P_i})]$

exponential form of Shannon's entropy possesses three key properties that support its application as an indicator of industrial variety (Straathof 2007). First, when all industries carry equal weight, the index equals the total number of industries. Second, the index is decomposable. Third, it is invariant to the addition or removal of industries with zero employment share. Variety, in this context, is understood as decomposable into two distinct components: inter-sectoral diversity, referred to as Unrelated Variety (UV), typically calculated at a higher level of aggregation (e.g., g-digit industry codes); and intra-sectoral diversity, or Related Variety (RV), measured at a more detailed level (e.g., i-digit codes). Higher values of UV and RV indicate lower industry concentration. From a theoretical perspective, Related Variety supports innovation through knowledge spillovers within similar industries, reflecting Jacobs' externalities. In contrast, Unrelated Variety captures the breadth of economic activities across sectors and serves—under portfolio theory—as a mechanism to buffer regions from sector-specific economic shocks, thereby enhancing labor market resilience. To improve readability and ensure consistency across sections, Table 1 provides a glossary of symbols and measures used throughout the paper.

Following Frenken et al. (2007), we organize economic activities in a tree with an overall root and two tiers. The root aggregates all activities. The upper tier (Level g) groups industries at the g-digit level, and the lower tier (Level i) contains the more granular i-digit categories (Fig. 1). This decomposition is classification-agnostic: it applies to any hierarchical taxonomy (e.g., NACE, ISIC, NAICS, SIC, or a bespoke scheme) provided groups are mutually exclusive and exhaustive and the ordering from fine to coarse is well defined.

Following Attaran (1986), the variety measure -VAR- will compare the regional employment distribution against a uniform employment distribution where the employment is equally distributed in all the industries. The measure of industrial Variety in each region is:²

² The choice of logarithm base in entropy-type variety measures is immaterial up to a constant rescaling. By the change-of-base identity, $\log_b x = \ln x / \ln b$, any entropy expression $H_b(p) = -\sum_i p_i \log_b(1/p_i)$ satisfies $H_2(p) = H_e(p) / \ln 2$ and $H_e(p) = (\ln 2) H_2(p)$. Consequently, a VAR index (and its decomposition into related and unrelated variety) computed with log base 2 is numerically equivalent to the same index computed with natural logs, up to the positive factor $1/\ln 2$. This preserves all economically relevant content—rankings across regions or sectors, comparative statics, and inference in regressions (coefficients and standard errors simply scale by the same constant)—while only changing the units (bits vs. nats).

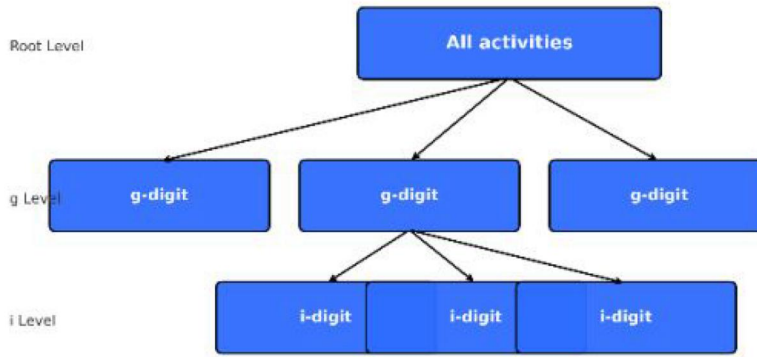


Figure 1. Two levels hierarchical structure.

$$\text{VAR}(P_1, P_2, \dots, P_n) = - \sum_{i=1}^N P_i \ln(P_i) \quad (2.1)$$

where $P_i = \frac{E_{ij}}{E_j}$, E_{ij} representing the number of employees in i -digit industry within region j , and E_j denoting total employment in region j . N is the total number of i -digit industries. The value of VAR lies within the interval $0 \leq \text{VAR} \leq \ln(N)$. The lower bound, $\text{VAR}=0$, represents maximum concentration and occurs when all employment is concentrated in a single industry (ie, $E_{ij} = E_j$ for one i and 0 for all others). The upper bound, $\text{AR} = \ln(N)$, reflects maximum dispersion and corresponds to a perfectly uniform distribution across all the industries. VAR depends solely on N , and the vector of employment shares $P = (P_1, P_2, \dots, P_n)$. It is a symmetric continuous function of P_i depending only on their relative magnitude and not on their order. If one industry, (say the n^{th}), is subdivided into two sub-industries with relative share q_1 and q_2 , the new measure of VAR equals the original VAR plus the conditional diversity within those industries (Hackbart and Anderson 1975). However, VAR does not identify the optimal distribution of employment, as different combinations of employment shares may yield the same level of variety. Using pre-defined hierarchical industry classification, the VAR index can be decomposed into Related Variety (RV) and Unrelated Variety (UV). Following Frenken et al., (2007) the i -digit industry can fall exclusively under g -digit industry, and P_g can be defined as:

$$P_g = \sum_{i \in g} P_i \quad (2.2)$$

where $P_g = \frac{E_{gj}}{E_j}$ is the employment at g -digit in region j on the total employment in region j . Using equation (2.2) equation (2.1) can be rewritten as:

$$\text{VAR} = \sum_{g=1}^G \sum_{i \in g} P_i \ln\left(\frac{1}{P_i}\right) \quad (2.3)$$

And rearranging³ the Variety in (2.1) becomes:

$$\text{VAR} = \text{RV} + \text{UV} \quad (2.4)$$

UV ranges from 0, -when all employment is concentrated in a single two-digit industry, to $\ln(G)$, when employment is evenly distributed across g -digit industry. RV ranges from 0 when each g -digit industry's employment is fully concentrated in one of its i -digit sub-industries- to $\ln(N) - \ln(G)$ when employment is uniformly distributed across all i -digit industries within each g -digit category (Castaldi et al., 2015, Fritsch and Kublina 2018).

³ A detailed derivation is provided in appendix 1.

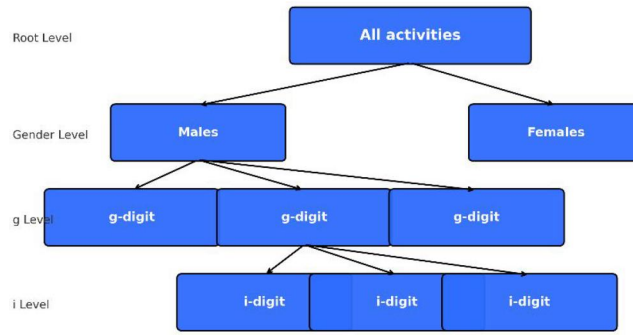


Figure 2. Three levels hierarchical structure.

3. Including gender into the analysis: the two-stage variety decomposition

Variety and its decomposition in Related and Unrelated Variety previously analyzed do not consider gender. Including gender in the analysis means including another dimension. Consequently, entropy decomposition will be based on a three-level hierarchical structure (Akita 2003, Wu et al., 2018). The hierarchical structure previously described needs to be modified, and Fig. 1 previously shown becomes:

Economic activities are organized into a three-tier hierarchy (Fig. 2): the gender level, the g level (g-digit industries), and the i level (i-digit industries). Total employment in region j , denoted E_j , is divided into K subgroups, with $K=1, 2$ where $K=1$ corresponds to the number of males and $K=2$ to the number of females employed in region j . Let $P_k = \frac{E_{kj}}{E_j}$ denote the share of females or males' workers in region j relative to the total employment i.e $P_{kg} = \frac{E_{kgj}}{E_j}$. Let P_{kg} denote the share of males or females employed in g-digit industry in region j , relative to total employment. Employment in each g-digit industry is then further distributed across its corresponding i-digit industries. Let P_{kgi} represent the share of male or female workers employed in i-digit industry in region j , again relative to total employment. According to the decomposition framework provided in appendix 2, the overall Variety measure can be expressed as follows:

$$\text{VAR} = \underbrace{\text{UV}_k + \text{UV}_{kg}}_{\text{UV}} + \underbrace{\text{RV}_F + \text{RV}_M}_{\text{RV}} \quad (3.1)$$

The first component UV_k captures gender-based diversity across the entire economy. It reflects disparities between men and women in terms of total employment distribution. UV_k ranges from 0 to $\ln(K)$ where $\text{UV}_k=0$ when $\frac{E_k}{E} = 1$ i.e when all employed individuals belong to the same gender, and it reaches its maximum in correspondence of $\ln(2)$ when men and women are equally distributed across the labor force. The measure is symmetric with respect to gender and reaches its peak under perfect parity. The second component, UV_{kg} , with $K=M; F$, denotes the Unrelated Variety across g-digit industries, considering the employment shares of males and females in each industry. It reflects inter-industry diversity by gender at an intermediate level of aggregation. The sum $\text{UV}_K + \text{UV}_{Kg}$ reproduces the overall Unrelated Variety (UV) as proposed by Frenken et al. (2007), now disaggregated by gender. Finally, RV_k with $K=M; F$ represents the Related Variety (RV) within i-digit industries, again computed separately for men and women. This component captures intra-industry diversity within each g-digit industry, based on the gender distribution at the most detailed level of classification.

4. Exploring related and unrelated variety measures

Building on the idea that gender can shape knowledge spillovers both within and across industries—and given that female and male employment shares differ systematically across sectors—we examine how Related Variety (RV) and Unrelated Variety (UV) respond to an increase in the female share. We consider three thought experiments: (i) changes in RV when female employment rises in a single i-digit industry i (holding all other cells constant); (ii) changes in UV at the g-digit level when female employment rises in a single g-digit industry; and (iii) changes in between-gender UV (UV_K) when female employment increases economy-wide. In addition, the appendix provides a tabular summary of the

methodological and interpretive implications of two broader employment-variation scenarios—proportional shocks and simultaneous multi-cell shifts—within the entropy-based RV/UV decomposition framework.

4.1 Question 1: How does RV change when the numbers of females increase in only one i-digit industry?

The results obtained in [appendix 3](#) highlight that if females' employment increases in only one i-digit industry, all the other things being equal, this change will impact the RV measure. Nevertheless, RV, under our decomposition, is the sum of RV_F and RV_M where RV_F captures the differences within industries at the i-digit level for females and RV_M captures the differences within industries at the i-digit level for males. Furthermore, our results highlight that RV_F and RV_M display different behavior. When female employment ($E_{F,i,j}$) increases in a single i-digit industry, RV_F initially increases if the original female shares in that industry ($s_{F,i,j}$) is below the equal distribution (ie, less than 50%). This suggests that introducing more females into a male-dominated industry is beneficial for knowledge spillovers among females. However, once the share reaches the equal distribution, further increases in female employment cause RV_F to decrease. This implies a decreasing return effect on Related Variety, where the diversification benefit plateaus and then diminishes as the share becomes highly skewed towards females ([Fig. 3.a](#)). Since male employment is held constant, the increase in female employment automatically leads to an increase in the industry's total employment, thereby reducing the proportional share of males. Consequently, RV_M shows a continuous decrease, following a hyperbolic trend ([Fig. 3.b](#)). The total RV (the sum of $RV_F + RV_M$) is thus determined by the relative magnitudes of the two opposing effects ([Fig. 3.c](#)). This leads to the conclusion that the total change in Related Variety is heavily dependent on the initial gender share in the industry. Gender has an impact on Related Variety, and this change is sensitive to the initial conditions (the current female share), exhibiting non-linear effects.

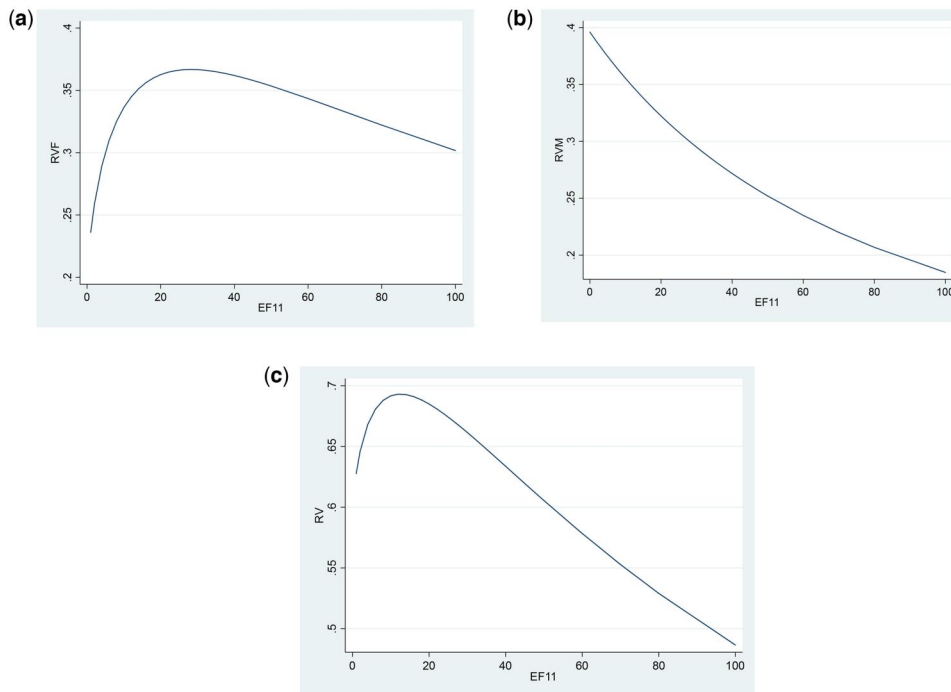


Figure 3. (a) Changes in RV_F when the numbers of females increase in only one 5-digit industry. (b) Changes in RV_M when the numbers of females increase in only one 5-digit industry. (c) Changes in RV when the numbers of females increase in only one 5-digit industry.

4.2 Question 2: How does UV_{Fg} changes when the numbers of females increase in only one g-digit industry?

The results obtained for UV_{kg} , as shown in [appendix 4](#), mirror those for Related Variety in their concavity. When female employment increases in a single g-digit industry, the female-specific measure UV_{Fg} rises as long as the distribution of female employment across industries moves closer to equipartition. Once this distribution becomes more concentrated in a single industry, UV_{Fg} begins to decrease. This reflects the well-known property of entropy, which attains its maximum under uniform distribution across all categories. In parallel UV_M , decreases monotonically as female employment expands while male employment is held constant, producing a hyperbolic-shaped decline. The total $UV = UV_F + UV_M$ therefore exhibits the same concave behavior identified for RV, confirming that gender composition significantly influences not only Related but also Unrelated Variety. Those results are depicted in [Fig. 4](#).

4.3 Question 3: How does UV_F changes when the numbers of females increase in the whole economy?

As shown in [appendix 5](#), UV_F has the same behaviors as RV_F and UV_{Fg} . It increases when $EF < EM$ and decreases when $EF > EM$. The relationship previously explored can be summarized as follows. Assume an increase of females' employment: $EF \uparrow \rightarrow UV_k$ changes. EM remains unchanged.

$UV_k \uparrow$	if	$EF < EM$
$UV_k \downarrow$	if	$EF > EM$

By hypothesis, EF will increase only the employment in one g-digit industry:

$$\Delta EF = \Delta E_{Fg:g=1}$$

$UV_{Mg} \downarrow$	if	$P_{Fg:g=1} = \frac{E_{Fg:g=1}}{E} < P_{Mg:g=1} = \frac{E_{Mg:g=1}}{E}$
$UV_{Fg} \uparrow$		
$UV_{Fg} \uparrow$ but with decreasing returns	if	$P_{Fg:g=1} = \frac{E_{Fg:g=1}}{E} > P_{Mg:g=1} = \frac{E_{Mg:g=1}}{E}$

By hypothesis $\Delta E_{Fg:g=1}$ will increase only the employment in one i-digit industry:

$$\Delta E_{Fg:g=1} = \Delta E_{Fi:i=1}$$

$RV_{Mg} \downarrow$	if	$P_{Fi:i=1} = \frac{E_{Fi:i=1}}{E} < P_{Mi:i=1} = \frac{E_{Mi:i=1}}{E}$
$RV_{Fg} \uparrow$		
$RV_{Fg} \uparrow$ but with decreasing returns	if	$P_{Fi:i=1} = \frac{E_{Fi:i=1}}{E} > P_{Mi:i=1} = \frac{E_{Mi:i=1}}{E}$

The UVF increases when the total female employment (EF) is less than the total male employment (EM) and decreases when EF exceeds EM. This general finding reinforces the central message: variety measures for a specific gender are maximized near the point of equal gender distribution ($EF \approx EM$), and the effect of employment growth is positive when starting from a minority position, but experiences a decreasing return once the equal distribution is surpassed.

These results show that gender composition acts as a moderating factor for knowledge spillovers (as captured by RV and UV). The largest boost to gender-specific variety is achieved when the gender that is in the minority increases its presence, driving the composition toward parity. Once parity is reached, further growth of that gender's share reduces the measure of variety for that group, indicating a loss of internal diversity among the group's members as the sector or economy becomes overwhelmingly dominated by them. This confirms two main conclusions: first, gender has a significant impact on both Related Variety and Unrelated Variety, second, the effect of a shock (employment increase) is dependent on the initial conditions (the pre-existing gender share).

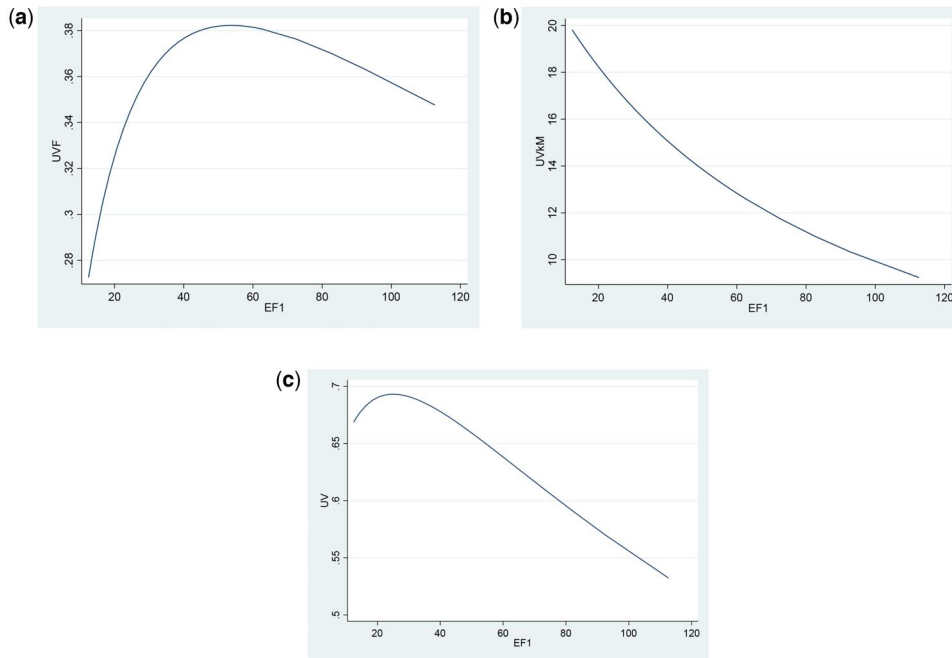


Figure 4. Changes in UV by gender when the numbers of females increase in only one g-digit industry. (a) Changes in UV_{Fg} when the numbers of females increase in only one 2-digit industry. (b) Changes in UVM_g when the numbers of males increase in only one 2-digit industry. (c) changes in UV when the numbers of females increase in only one 2-digit industry.

5. Empirical evidence

The theoretical results should be validated through empirical analysis. Data on employment by sector of economic activity is provided by Istat through the Labour Force Survey (LFS). The analysis is carried out at the regional level, and the list of Italian regions is reported in [Table 2.A.](#) in the [Appendix Information](#) at the i-digit ATECO level is not publicly available but can be accessed upon request. For the empirical analysis, we therefore rely on the LFS dataset, which classifies economic activities into twelve categories (CAT12), used here as the i-level. In our case, the g-level corresponds to CAT5 in the same dataset, as reported in [Table 3.A](#) in the [appendix](#). The observation period spans from the first quarter of 2008 to the first quarter of 2025. However, following the entry into force of Regulation (EU) 2019/1700 on 1 January 2021, the LFS no longer includes regional identifiers. Consequently, our dataset covers the period 2008–2020. The data are quarterly, while annual figures are computed as the average for the four quarters. The list of Italian regions is reported in [Table 1.A](#). For confidentiality reasons, data for Piemonte and Valle d’Aosta are aggregated and reported as a single unit. [Figure 5](#) shows the share of employment by region, category, and gender. As can be seen, the share of male employment varies not only across sectors but also across regions within the same category.

This heterogeneity is also reflected in terms of Related and Unrelated Variety, as shown in [Figs. 6.a](#) and [6.b](#).

The theoretical analysis developed in the previous section has shown that the RV and UV measures so far proposed in the literature, when disaggregated by gender, respond differently to changes in male and female employment. They exhibit increasing returns to scale when women are underrepresented in a given industry but decreasing returns when they are overrepresented. The same holds for men. To empirically test these measures, we relate them to regional Value Added (VA) for the period 2008–2020. The estimated relationship is as follows:

$$\log(VA_{jt}) = \alpha_j + \beta_1 RVM_{jt} + \beta_2 RVF_{jt} + \beta_3 UVM_{jt} + \beta_4 UVF_{jt} + \beta_5 UV_{Kjt} + \varepsilon_{jt} \quad (5.1)$$

Where j denotes the region and t the time. [Equation \(5.1\)](#) is estimated using a fixed-effects panel model. Because the covariates are correlated, we evaluated alternative covariate sets and selected the specification with the lowest multicollinearity.

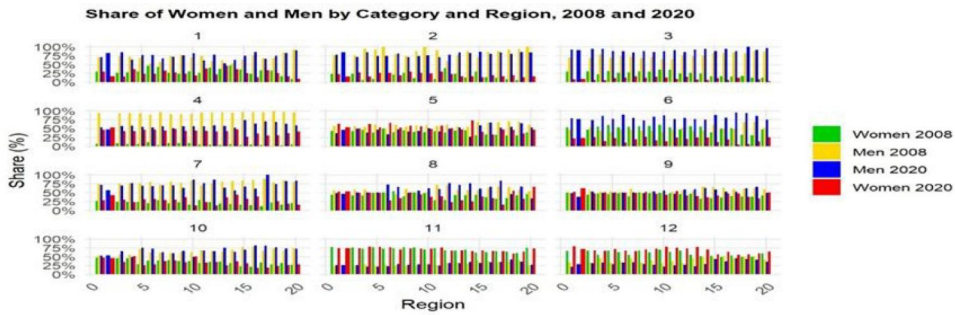


Figure 5. Share of employment by region, category, and gender; 2008 and 2020.

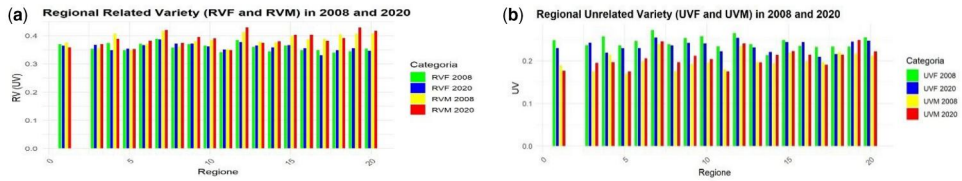


Figure 6. Regional related and unrelated variety by gender, 2008 and 2020. (a) Related variety by region and gender 2008 and 2020. (b) Unrelated variety by region and gender 2008 and 2020.

Table 2. Estimation results.

Dependent Variable $\log(VA)$	
RVF	1.744*** (0.0009)
UVM	-0.3117 (0.486)
UVF	2.2983*** (4.19E-09)
UVK	8.0694*** (9.70E-06)

P value in parenthesis. Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $p < 0.1$.

$$\log(VA_{jt}) = \alpha_j + \beta_2 RVF_{jt} + \beta_3 UVM_{jt} + \beta_4 UVF_{jt} + \beta_5 UVK_{jt} + \varepsilon_{jt} \quad (5.2)$$

Table 2 reports the estimation results with $\log(VA)$ as the dependent variable. The coefficients indicate that gender-sensitive measures of Related and Unrelated Variety may be relevant to explaining regional differences in value-added growth. The coefficient of RVF (Related Variety—Female) is positive and highly significant, suggesting that intra-industry diversification associated with a stronger female presence may be linked to better regional economic performance. This pattern is broadly consistent with the notion that women’s participation in industries where they remain underrepresented could enhance the scope for cognitive spillovers and innovation, although this mechanism cannot be formally verified within the current framework. By contrast, UVM (Unrelated Variety—Male) is negative and not statistically significant, implying that diversification across industries driven mainly by male employment does not seem to contribute meaningfully to regional growth. The coefficient of UVF (Unrelated Variety—Female) is positive and significant, indicating that a wider distribution of female employment across industries may be associated with improved regional outcomes. This association might reflect portfolio effects or greater resilience to asymmetric shocks, yet these interpretations should be regarded as indicative rather than conclusive. Finally, UVK (Unrelated Variety by Gender) displays a large positive and significant coefficient, which may suggest that overall gender diversity in regional labour markets is positively associated with value-added growth. Such diversity could support complementarities and absorptive capacity, thereby enhancing regional adaptability, although causality cannot be firmly established. Taken together, these results provide preliminary evidence that gender composition interacts with the dynamics of related and unrelated variety in shaping regional performance.

6. Discussion and policy implications

The Evolutionary Economic Geography (EEG) framework has long viewed knowledge spillovers as a central driver of regional growth, yet it has remained largely gender-neutral. While recent contributions in economic geography have begun to acknowledge gendered patterns in innovation and labor markets, widely used metrics—most notably Related and Unrelated Variety (RV and UV) as developed by [Frenken et al. \(2007\)](#)—still treat the workforce as homogeneous. By contrast, a substantial body of social-science research has shown that women and men differ systematically in innovation styles, inter-industry collaboration, knowledge-sharing practices, occupational choices, and workplace networks. These differences shape how knowledge circulates within and across industries, especially in contexts where gender segregation is pronounced. This paper extends the RV/UV framework with a gender-sensitive decomposition that disaggregates variety by gender and distinguishes intra- from inter-industry components. Theoretical results reveal that variety responds asymmetrically to changes in gender composition: when women are underrepresented, increases in female employment raise both related and unrelated variety; once parity is reached or exceeded, marginal gains diminish. These dynamics imply that the scope for knowledge recombination is greatest in contexts of strong gender segregation, where incremental participation by the underrepresented group generates disproportionate spillover effects. Empirically, using quarterly Istat Labour Force Survey data (2008–2020), we show that female-specific measures of variety are strongly associated with regional value added. In particular, both RVF and UVF display positive and highly significant coefficients, while overall gender diversity (UVK) is also strongly positive and UVM remains statistically insignificant. These patterns are broadly consistent with the view that greater female participation in male-dominated industries may be associated with enhanced cognitive spillovers and innovation potential, although the evidence should be interpreted with caution given the correlational nature of the analysis. The policy implications are clear. Rather than reinforcing female concentration in already female-dominated sectors, interventions should prioritize “low-participation, high-relatedness” niches. Concrete measures include apprenticeships and scholarships in adjacent STEM and tradable-service domains, procurement and R&D incentives tied to gender-diversity conditions, and childcare and flexible-work reforms that unlock participation where spillovers are most likely. Gender balance is therefore not only a matter of equity; it also represents a productive strategy for fostering regional resilience and long-term growth. These findings resonate with current European and national debates on gender and innovation, aligning with the European Union’s commitment to gender mainstreaming within Smart Specialization Strategies (S3), where inclusivity is viewed as a core condition for effective innovation ecosystems. They further support EU initiatives promoting women’s participation in STEM and digital fields, such as Horizon Europe’s gender equality requirements and national action plans. Incorporating gender-sensitive dimensions of related and unrelated variety into development strategies may thus enhance both the fairness and the effectiveness of innovation policy. By explicitly linking gender equality to productivity outcomes, this paper suggests that encouraging women’s participation in high-relatedness, male-dominated sectors is not only socially just but also economically strategic for sustainable regional development.

We acknowledge some limitations. The proposed measures have not yet been systematically benchmarked against established segregation indices ([Duncan and Duncan 1955](#); [Karmel and MacLachlan 1988](#)) and remain complementary to network-based relatedness metrics ([Hidalgo et al. 2007](#); [Neffke and Henning 2013](#)), which, while powerful, are still gender-blind. Future research should combine our entropy-based indicators with these benchmarks and explore multilayer network approaches to capture gendered recombination dynamics more directly. Advancing along these lines will sharpen identification, strengthen external validity, and move EEG toward a framework where gender is recognized as a structural dimension of innovation and resilience.

Conflict of interest statement. None declared.

Funding

None declared.

Appendix

Appendix 1: Related and Unrelated Variety using a one-stage decomposition

Multiplying for $\frac{P_g}{P_g}$, equation (2.1) becomes:

$$\text{VAR} = \sum_{g=1}^G \left[\sum_{i \in g} \frac{P_g}{P_g} * P_i \left(\ln \frac{P_g}{P_g} \frac{1}{P_i} \right) \right]$$

Applying the log properties, and rearranging:

$$\text{VAR} = \sum_{g=1}^G \left[\sum_{i \in g} \frac{P_g}{P_g} * P_i \left(\ln \left(\frac{P_g}{P_i} \right) + \ln \left(\frac{1}{P_g} \right) \right) \right] = \sum_{g=1}^G \left[\sum_{i \in g} \frac{P_i}{P_g} * P_g \ln \left(\frac{P_g}{P_i} \right) + \sum_{i \in g} P_i \ln \left(\frac{1}{P_g} \right) \right] \quad (\text{A1.1})$$

Using equation (2.2), equation (A1.3) can be re-written as:

$$\text{VAR} = \sum_{g=1}^G P_g \left[\sum_{i \in g} \frac{P_i}{P_g} \ln \left(\frac{P_g}{P_i} \right) \right] + \left[\sum_{g=1}^G P_g \ln \left(\frac{1}{P_g} \right) \right] \quad (\text{A1.2})$$

Defining:

$$H_g = \left[\sum_{i \in g} \frac{P_i}{P_g} \ln \left(\frac{P_g}{P_i} \right) \right] \quad (\text{A1.3})$$

$$\text{RV} = \sum_{g=1}^G P_g H_g \quad (\text{A1.4})$$

$$\text{UV} = \left[\sum_{g=1}^G P_g \ln \frac{1}{P_g} \right] \quad (\text{A1.5})$$

Appendix 2: Related and Unrelated Variety using a two-stage decomposition

By this new classification equation (A1.1) becomes:

$$P_{kg} = \sum_{i \in g} P_{ki} \quad (\text{A2.1})$$

and equation (A1.2) becomes:

$$\text{VAR} = \sum_{k=1}^2 \sum_{g=1}^G \sum_{i \in g} P_{ki} \ln \left(\frac{1}{P_{ki}} \right) \quad (\text{A2.2})$$

Multiplying for $\frac{P_{kg}}{P_{kg}}$ equation (A2.2) can be rewritten as:

$$\begin{aligned} \text{VAR} &= \sum_{k=1}^2 \sum_{g=1}^G \sum_{i \in g} P_{ki} \frac{P_{kg}}{P_{kg}} \ln \left(\frac{1}{P_{ki}} \frac{P_{kg}}{P_{kg}} \right) = \sum_{k=1}^2 \sum_{g=1}^G \sum_{i \in g} P_{ki} \frac{P_{kg}}{P_{kg}} \left[\ln \left(\frac{1}{P_{ki}} \right) + \ln \left(\frac{P_{kg}}{P_{kg}} \right) \right] = \text{VAR} \\ &= \sum_{k=1}^2 \sum_{g=1}^G \left[\sum_{i \in g} P_{ki} \ln \left(\frac{1}{P_{ki}} \right) + \sum_{i \in g} P_{kg} \frac{P_{ki}}{P_{kg}} \ln \left(\frac{P_{kg}}{P_{ki}} \right) \right] \end{aligned} \quad (\text{A2.3})$$

Using equation (A2.1) equation (A2.3) can be rewritten as:

$$\text{VAR} = \sum_{k=1}^2 \left(\sum_{g=1}^G P_{kg} \ln \left(\frac{1}{P_{kg}} \right) + \sum_{g=1}^G P_{kg} \sum_{i \in g} \frac{P_{ki}}{P_{kg}} \ln \left(\frac{P_{kg}}{P_{ki}} \right) \right) \quad (\text{A2.4})$$

Equation (A1.5) now becomes:

$$H_{kg} = \left[\sum_{i \in g} \frac{P_{ki}}{P_{kg}} \ln \left(\frac{P_{kg}}{P_{ki}} \right) \right] \quad (\text{A2.5})$$

where H_{kg} represents the Variety by gender within groups with $\sum_{k=1}^2 H_{kg} = H_g$ as in equation (A1.5). Using equation (A2.5) the equation (A2.4) can be rewritten as:

$$\text{VAR} = \sum_{k=1}^2 \left(\sum_{g=1}^G P_{kg} \ln \left(\frac{1}{P_{kg}} \right) + \sum_{g=1}^G P_{kg} H_{kg} \right) = \sum_{k=1}^2 \sum_{g=1}^G P_{kg} \ln \left(\frac{1}{P_{kg}} \right) + \sum_{k=1}^2 \sum_{g=1}^G P_{kg} H_{kg} \quad (\text{A2.6})$$

Equation (A2.6) is the sum of two components. The second one:

$$\text{RV} = \sum_{k=1}^2 \sum_{g=1}^G P_{kg} H_{kg} = \sum_{k=1}^2 \text{RV}_k$$

ie the sum of Related Variety by gender, while the first term

$$\text{UV} = \sum_{k=1}^2 \sum_{g=1}^G P_{kg} \ln \left(\frac{1}{P_{kg}} \right) = \sum_{k=1}^2 \text{UV}_k \quad (\text{A2.7})$$

i.e the sum of Unrelated Variety by gender.

Equation (A1.8) becomes:

$$\text{VAR} = \sum_{k=1}^2 \text{UV}_k + \sum_{k=1}^2 \text{RV}_k \quad (\text{A2.8})$$

Where $\sum_{k=1}^2 \text{UV}_k = \text{UV}$ and $\sum_{k=1}^2 \text{RV}_k = \text{RV}$.

The equation (A2.8) is equal to equation (A1.8) used by Frenken et al. (2007). The novelty is represented by the decomposition in females and males. This new measure allows us to explore if and how gender affects the innovation process.

The UV in equation (A2.7) can be furtherly decomposed. Indicating

$$P_k = \sum_{g \in k} P_{kg} \quad (\text{A2.9})$$

and multiplying equation (A2.7) by $\frac{P_{kg}}{P_{kg}}$ we obtain:

$$\text{UV} = \sum_{k=1}^2 \left[\sum_{g \in k} \frac{P_k}{P_k} * P_{kg} \ln \left(\frac{P_k}{P_k} \frac{1}{P_{kg}} \right) \right] = \sum_{k=1}^2 \left[\sum_{g \in k} \frac{P_k}{P_k} * P_{kg} \ln \left(\frac{P_k}{P_{kg}} \right) + \sum_{g \in k} \frac{P_k}{P_k} * P_{kg} \ln \left(\frac{1}{P_k} \right) \right]$$

Using equation (A2.1) we obtain:

$$\text{UV} = \sum_{k=1}^2 P_k \sum_{g \in k} \frac{P_{kg}}{P_k} \ln \left(\frac{P_k}{P_{kg}} \right) + \sum_{k=1}^2 P_k \ln \left(\frac{1}{P_k} \right) \quad (\text{A2.10})$$

and indicating H_k as:

$$H_k = \sum_{g \in k} \frac{P_{kg}}{P_k} \ln \left(\frac{P_k}{P_{kg}} \right)$$

Equation (A2.10) becomes:

$$UV = \sum_{k=1}^2 P_k H_k + \sum_{k=1}^2 P_k \ln \left(\frac{1}{P_k} \right)$$

where H_k represents the unrelated Variety by gender between industries at a g-digit level. It varies from 0 to $\ln(G) - \ln(K)$. $H_k=0$ when $P_k = P_{kg}$ i.e all the employment of a given gender belongs to the same g-digit industry. When $H_k = \ln(G) - \ln(K)$ genders are equidistributed between g-digit industries. The first term:

$$UV_{kg} = \sum_{k=1}^2 P_k H_k$$

represents the weighted unrelated Variety by gender between industries at a g-digit level while the second term:

$$UV_k = \sum_{k=1}^2 P_k \ln \left(\frac{1}{P_k} \right)$$

represents the diversity by gender between all the economic activities and it captures the disparities between females and males in the whole economy. UV_k varies from 0 to $\ln(2)$. $UV_k = 0$ when $\frac{E_k}{E} = 1$ i.e when all employed belong to the same gender. When $UV_k = \ln(2)$ females and males are equidistributed between economic activities. The term UV_k is symmetric to the females' (males') and it reaches its maximum in correspondence of $\ln(2)$. The sum of UV_k and UV_{kg} gives the UV as in [Frenken et al. \(2007\)](#). Finally, [equation \(A2.8\)](#) can be rewritten as:

$$VAR = \underbrace{UV_{kg} + UV_k}_{UV} + \underbrace{RV_F + RV_M}_{RV} \quad (A2.11)$$

Appendix 3: RV changes when the numbers of females increase only in one i-digit industry.

In accordance with the previous decomposition the RV is given by:

$$RV = \sum_{k=1}^2 \sum_{g=1}^G P_{kg} H_{kg} = \sum_{k=1}^2 RV_k \quad (A3.1)$$

Where H_{kg} is given by [equation \(A2.5\)](#). RV can be rewritten as:

$$RV = \sum_{g=1}^G P_{Fg} H_{Fg} + \sum_{g=1}^G P_{Mg} H_{Mg} = \sum_{g=1}^G P_{Fg} \left[\sum_{i \in g} \frac{P_{Fi}}{P_{Fg}} \ln \left(\frac{P_{Fg}}{P_{Fi}} \right) \right] + \sum_{g=1}^G P_{Mg} \left[\sum_{i \in g} \frac{P_{Mi}}{P_{Mg}} \ln \left(\frac{P_{Mg}}{P_{Mi}} \right) \right] \quad (A3.2)$$

Where:

$$P_{Fi} = \frac{E_{Fi}}{E}; P_{Fg} = \frac{E_{Fg}}{E}; P_{Mi} = \frac{E_{Mi}}{E}; P_{Mg} = \frac{E_{Mg}}{E};$$

[Equation \(A3.2\)](#) can be rewritten as:

$$RV = \underbrace{\sum_{g=1}^G \frac{E_{Fg}}{E} \left[\sum_{i \in g} \frac{E_{Fi}}{E_{Fg}} \ln \left(\frac{E_{Fg}}{E_{Fi}} \right) \right]}_{RV_F} + \underbrace{\sum_{g=1}^G \frac{E_{Mg}}{E} \left[\sum_{i \in g} \frac{E_{Mi}}{E_{Mg}} \ln \left(\frac{E_{Mg}}{E_{Mi}} \right) \right]}_{RV_M} \quad (A3.3)$$

Remembering that: $E = E_F + E_M$; $E_F = \sum_{i=1}^N E_{Fi}$; $E_{Fg} = \sum_{i \in g} E_{Fi}$ equation (A3.3) can be rewritten as:

$$RV_F = \frac{1}{(E_M + E_F)} \left\{ \sum_{g=1}^G \left[\sum_{i \in g} E_{Fi} \ln \left(\frac{E_{Fg}}{E_{Fi}} \right) \right] \right\} \quad (A3.4.1)$$

$$RV_M = \frac{1}{(E_M + E_F)} \left\{ \sum_{g=1}^G \left[\sum_{i \in g} E_{Mi} \ln \left(\frac{E_{Mg}}{E_{Mi}} \right) \right] \right\} \quad (A3.4.2)$$

To investigate the RV changes when the females increase in only one i -digit industry the RV derivative with respect to the females' change needs to be calculated. We suppose that females' employment increase while males' employment remains unchanged. Consequently, the term $\sum_{g=1}^G \left[\sum_{i \in g} E_{Mi} \ln \left(\frac{\sum_{i \in g} E_{Mi}}{E_{Mi}} \right) \right] = \text{cost}$. Our aim is to investigate the changes in RV_F when the females' employment changes only in one of the i industries belonging to g with $g=1$. Indicating with x the females' employment in industry $i=1$ (E_{F1}), equations (A3.4.1) (A3.4.2) can be rewritten as:

$$RV_F = \frac{1}{x + \sum_{i=2}^N E_{Fi} + E_M} \left\{ \sum_{g=1}^G \left[\sum_{i \in g} E_{Fi} \ln \left(\frac{\sum_{i \in g} E_{Fi}}{E_{Fi}} \right) \right] \right\} \quad (A3.5.1)$$

$$RV_M = \frac{\text{cost}}{x + \sum_{i=2}^N E_{Fi} + E_M} \quad (A3.5.2)$$

The terms $\sum_{i=2}^N E_{Fi} + E_M$ in equation (4.4.1) are constant with respect to the change of E_{F1} .

$$c_1 = \sum_{i=2}^N E_{Fi} + E_M$$

$$RV_F = \frac{1}{x + c_1} \left\{ \sum_{g=1}^G \left[\sum_{i \in g} E_{Fi} \ln \left(\frac{\sum_{i \in g} E_{Fi}}{E_{Fi}} \right) \right] \right\} \quad (A3.5.1 \text{ bis})$$

By hypothesis industry $i=1$ can fall exclusively in one g . Suppose that $i=1 \in g=1$. The content of the square brackets in equation A.1 can be rewritten as:

$$\sum_{g=1}^G \left[\sum_{i \in g} E_{Fi} \ln \left(\frac{\sum_{i \in g} E_{Fi}}{E_{Fi}} \right) \right] = \sum_{g=2}^G \left[\sum_{i \in g; g \neq 1} E_{Fi} \ln \left(\frac{\sum_{i \in g; g \neq 1} E_{Fi}}{E_{Fi}} \right) \right] + \left[\sum_{i \in g; g=1} E_{Fi} \ln \left(\frac{\sum_{i \in g; g=1} E_{Fi}}{E_{Fi}} \right) \right]$$

Moreover, term $\sum_{g=2}^G \left[\sum_{i \in g; g \neq 1} E_{Fi} \ln \left(\frac{E_{Fg}}{E_{Fi}} \right) \right]$ is constant with respect to E_{F1} , so:

$$c_2 = \sum_{g=2}^G \left[\sum_{i \in g; g \neq 1} E_{Fi} \ln \left(\frac{\sum_{i \in g} E_{Fi}}{E_{Fi}} \right) \right]$$

$$RV_F = \frac{1}{x + c_1} \left\{ c_2 + \left[\sum_{i \in g; g=1} E_{Fi} \ln \left(\frac{\sum_{i \in g; g=1} E_{Fi}}{E_{Fi}} \right) \right] \right\}$$

The term can be rewritten as:

$$\left[\sum_{i \in g; g=1} E_{Fi} \ln \left(\frac{\sum_{i \in g; g=1} E_{Fi}}{E_{Fi}} \right) \right] = x * \ln \left(\frac{x + \sum_{i \neq 1 \in g; g=1} E_{Fi}}{x} \right) + \sum_{i \neq 1 \in g; g=1} E_{Fi} * \ln \left(\frac{x + \sum_{i \neq 1 \in g; g=1} E_{Fi}}{E_{Fi}} \right)$$

Where $\sum_{i \neq 1 \in g; g=1} E_{Fi}$ is constant with respect to x :

$$c_3 = \sum_{i \neq 1 \in g; g=1} E_{Fi}$$

Consequently:

$$RV_F = \frac{1}{x + c_1} \left\{ c_2 + \left[x * \ln \left(\frac{x + c_3}{x} \right) + c_3 \ln \left(\frac{x + c_3}{E_{Fi}} \right) \right] \right\}$$

Applying the log properties to the term:

$$c_3 \ln \left(\frac{x + c_3}{E_{Fi}} \right)$$

We obtain:

$$c_3 [\ln(x + c_3) - \ln(E_{Fi})] = c_3 \ln(x + c_3) - c_3 \ln(E_{Fi})$$

Moreover, the term $c_3 \ln(E_{Fi})$ is constant with respect to x . We indicate it as $c_4 = c_3 \ln(E_{Fi})$; Finally,

$$c_3 \ln \left(\frac{x + c_3}{E_{Fi}} \right) = c_3 \ln(x + c_3) - c_4$$

RV_F is equal to:

$$RV_F = \frac{1}{x + c_1} \left\{ c_2 + \left[x * \ln \left(\frac{x + c_3}{x} \right) + c_3 \ln(x + c_3) - c_4 \right] \right\}$$

$$RV_F = \frac{1}{x + \sum_{i=2}^N E_{Fi} + E_M} \left\{ \sum_{g=2}^G \left[\sum_{i \in g; g \neq 1} E_{Fi} \ln \left(\frac{\sum_{i \in g} E_{Fi}}{E_{Fi}} \right) \right] + \left[x * \ln \left(\frac{x + \sum_{i \neq 1 \in g; g=1} E_{Fi}}{x} \right) \right. \right.$$

$$\left. \left. + \left[\ln \left(x + \sum_{i \neq 1 \in g; g=1} E_{Fi} \right) \right] \sum_{i \neq 1 \in g; g=1} E_{Fi} - \left(\sum_{i \neq 1 \in g; g=1} E_{Fi} \right) \ln(E_{Fi}) \right] \right\} \tag{A3.6}$$

To study how RV changes when E_{F1} changes, the first derivative needs to be investigated. RV_M derivatives respect to E_{F1} is:

$$\frac{\partial RV_M}{\partial E_{F1}} = - \frac{\sum_{g=1}^G \left[\sum_{i \in g} E_{Mi} \ln \left(\frac{\sum_{i \in g} E_{Mi}}{E_{Mi}} \right) \right]}{\left(x + \sum_{i=2}^N E_{Fi} + E_M \right)^2}$$

$$\lim_{x \rightarrow \infty} \frac{\partial RV_M}{\partial E_{F1}} = 0$$

$$\text{for } x = 0 \text{ RV}_M = - \frac{\sum_{g=1}^G \left[\sum_{i \in g} E_{Mi} \ln \left(\frac{\sum_{i \in g} E_{Mi}}{E_{Mi}} \right) \right]}{\left(\sum_{i=2}^N E_{Fi} + E_M \right)^2}$$

RV_F derivatives respect to E_{F1} is:

$$\frac{\partial \text{RV}_F}{\partial E_{F1}} = \frac{\ln \left(\frac{x + \sum_{i \neq 1; g \neq 1} E_{Fi}}{x} \right) \left(\sum_{i=2}^N E_{Fi} + E_M \right) - \ln \left(x + \sum_{i=2}^N E_{Fi} + E_M \right) \left(\sum_{i=2}^N E_{Fi} + E_M \right) - \sum_{g=2}^G \left[\sum_{i \in g; g \neq 1} E_{Fi} \ln \left(\frac{\sum_{i \in g} E_{Fi}}{E_{Fi}} \right) \right] + \sum_{i \neq 1; g=1} E_{Fi} \ln(E_{Fi})}{\left(x + \sum_{i=2}^N E_{Fi} + E_M \right)^2}$$

$$\lim_{x \rightarrow \infty} \frac{\partial \text{RV}_F}{\partial E_{F1}} = 0$$

$$\lim_{x \rightarrow 0} \frac{\partial \text{RV}_F}{\partial E_{F1}} = \infty$$

$\frac{\partial \text{RV}_F}{\partial E_{F1}} = 0$ for $x = x^*$.

For $0 < x < x^*$ $\frac{\partial \text{RV}_F}{\partial E_{F1}}$ increases while for $x^* < x < \infty$ $\frac{\partial \text{RV}_F}{\partial E_{F1}}$ decreases. x^* represent the point in which females and males are equidistributed between gender.

Appendix 4: UV_{kg} changes when the numbers of females increase in only one g-digit industry

In accordance with the previous results UV_{kg} is given by:

$$\text{UV}_{kg} = \sum_{k=1}^2 P_k H_k$$

$$H_k = \sum_{g=1}^G \frac{P_{kg}}{P_k} \ln \left(\frac{P_k}{P_{kg}} \right)$$

$$\text{UV}_{kg} = \underbrace{P_F * \sum_{g=1}^G \frac{P_{Fg}}{P_F} \ln \left(\frac{P_F}{P_{Fg}} \right)}_{\text{UV}_{Fg}} + \underbrace{P_M * \sum_{g=1}^G \frac{P_{Mg}}{P_M} \ln \left(\frac{P_M}{P_{Mg}} \right)}_{\text{UV}_{Mg}}$$

Where $P_F = \frac{E_F}{E}$, $E = E_F + E_M$, $E_F = \sum_{g=1}^G E_{Fg}$. UV_{kg} can be rewritten as:

$$\text{UV}_{kg} = \underbrace{\frac{E_F}{\sum_{g=1}^G E_{Fg} + E_M} * \frac{1}{E_F} \sum_{g=1}^G E_{Fg} \ln \left(\frac{E_F}{E_{Fg}} \right)}_{\text{UV}_{Fg}} + \underbrace{\frac{E_M}{\sum_{g=1}^G E_{Fg} + E_M} * \frac{1}{E_M} \sum_{g=1}^G E_{Mg} \ln \left(\frac{E_M}{E_{Mg}} \right)}_{\text{UV}_{Mg}} \quad (\text{A4.1})$$

The term $\sum_{g=1}^G E_{Mg} \ln \left(\frac{E_M}{E_{Mg}} \right)$ is constant with respect to E_{F1}, consequently, $c_1 = \sum_{g=1}^G E_{Mg} \ln \left(\frac{E_M}{E_{Mg}} \right)$, furthermore, the term $\sum_{g=1}^G E_{Fg}$ can be rewritten as $E_{F1} + \sum_{g=2}^G E_{Fg}$. Equation (A4.1) becomes:

$$UV_{kg} = \frac{E_{F1} * \ln\left(\frac{E_F}{E_{F1}}\right) + \sum_{g=2}^G E_{Fg} \ln\left(\frac{E_F}{E_{Fg}}\right)}{\underbrace{E_{F1} + \sum_{g=2}^G E_{Fg} + E_M}_{UV_{Fg}}} + \frac{c_1}{\underbrace{E_{F1} + \sum_{g=2}^G E_{Fg} + E_M}_{UV_{Mg}}} \tag{A4.2}$$

The terms $\sum_{g=2}^G E_{Fg} \ln\left(\frac{E_F}{E_{Fg}}\right)$ and $\sum_{g=2}^G E_{Fg} + E_M$ are constant with respect to E_{F1} .

$$c_2 = \sum_{g=2}^G E_{Fg} + E_M; c_3 = \sum_{g=2}^G E_{Fg} \ln\left(\frac{E_F}{E_{Fg}}\right)$$

Equation (A4.2) becomes:

$$UV_{kg} = \frac{E_{F1} * \ln\left(\frac{E_F}{E_{F1}}\right) + c_3}{\underbrace{E_{F1} + c_2}_{UV_{Fg}}} + \frac{c_1}{\underbrace{E_{F1} + c_2}_{UV_{Mg}}} \tag{A4.3}$$

$$\frac{\partial UV_{Mg}}{\partial EF_1} = \frac{-c_1}{(c_2 + E_{F1})^2}$$

$$\lim_{x \rightarrow \infty} \frac{\partial UV_{Mg}}{\partial EF_1} = 0$$

$$\frac{\partial UV_{Mg}}{\partial EF_1} = \frac{-c_1}{(c_2)^2} \text{ when } EF_1 = 0$$

$$\frac{\partial UV_{Fg}}{\partial EF_g} = \frac{-c_3 + c_2 * \ln\left(\frac{E_F}{E_{F1}}\right) - E_{F1} - c_2}{(E_{F1} - c_2)^2}$$

$$\lim_{x \rightarrow \infty} \frac{\partial UV_{Fg}}{\partial EF_1} = 0$$

$$\lim_{x \rightarrow 0} \frac{\partial UV_{Fg}}{\partial EF_1} = \infty$$

Finally, $\frac{\partial UV_{Fg}}{\partial EF_1} = 0$ for $x = x^*$.

Appendix 5: UV_F changes when the numbers of females increase in the whole economy

$$UV_k = \sum_{k=1}^2 P_k \ln\left(\frac{1}{P_k}\right)$$

$$UV_k = \frac{E_M}{E} \ln\left(\frac{E}{E_M}\right) + \frac{E_F}{E} \ln\left(\frac{E}{E_F}\right)$$

$$E = E_M + E_F$$

$$UV_k = \frac{E_M}{E_M + E_F} \ln\left(\frac{E_M + E_F}{E_M}\right) + \frac{E_F}{E_M + E_F} \ln\left(\frac{E_M + E_F}{E_F}\right)$$

$$E_M = c; E_F = x$$

$$UV_k = \frac{c}{c+x} \ln\left(\frac{c+x}{c}\right) + \frac{x}{c+x} \ln\left(\frac{c+x}{x}\right)$$

$$\frac{\partial UV_k}{\partial x} = -\frac{c \left[\ln\left(\frac{c+x}{c}\right) - \ln\left(\frac{c+x}{x}\right) \right]}{(c+x)^2}$$

$$\lim_{x \rightarrow \infty} \frac{\partial UV_k}{\partial EF} = 0$$

$$\lim_{x \rightarrow 0} \frac{\partial UV_k}{\partial EF} = \infty$$

Finally, $\frac{\partial UV_k}{\partial EF} = 0$ when $EF = EM$ i.e when females and males are equi-distributed in the whole economy.

Appendix 6: Details on Experimental Results and Entropy Decompositions

This appendix presents a tabular summary of the methodological and interpretive implications of two employment-variation scenarios—proportional shocks and simultaneous multi-cell shifts—within the entropy-based decomposition of variety.

Table 1.A Implications of employment variation scenarios for entropy-based variety decomposition.

Scenario	Where EF increases	RVF (within g-digit)	UVF (across g-digit)	RVM/UVM (men fixed)	UVK	Notes
Targeted push into underrepresented cells	i-digit with $S_{F g(i)} < \text{equiproportion}$; g-digit with $S_F(g) < \text{equiproportion}$	↑	↑	↓	=	Stronger effect when cells were far below equiproportion and the increases are large.
Balanced mix (under + over)	Some cells below and some above equiproportion	± (attenuated)	± (attenuated)	↓ (on average)	=	First-order effects add up across cells; positives and negatives may almost cancel.
Push into already dominant (overrepresented) cells	i-digit/g-digit with female shares > equiproportion	↓	↓	↓	=	Greater concentration reduces within-gender and between-group variety.
Proportional shock (same % in all cells)	All i-digit/g-digit	=	=	=	$1/U =$ (depends on EF/EM)	Within-gender shares do not change; only the between-gender component moves.
Price-taking regional economy (both genders adjust)	$\Delta E_F \neq 0$ and $\Delta E_M \neq 0$	Same signs as A-C if female growth is relatively stronger where F is underrepresented	Ditto	Depends on relative changes	Depends	If male growth offsets female gains, RVF/UVF effects attenuate or can reverse.
Multi-cell, additivity rule	Multiple cells simultaneously	Sum of per-cell contributions	Sum of per-group contributions	Sum (typically ↓ with men fixed)	Sum (if relevant)	Entropy decomposition is additively separable: first-order effect is the sum of single-cell effects.

Appendix 7: Data Classifications: Italian Regions and CAT-12 Sectors

Table 2A. Italian regions.

1	Piemonte	11	Marche
2	Valle d'Aosta	12	Lazio
3	Lombardia	13	Abruzzo
4	Trentino Alto Adige	14	Molise
5	Veneto	15	Campania
6	Friuli Venezia Giulia	16	Puglia
7	Liguria	17	Basilicata
8	Emilia Romagna	18	Calabria
9	Toscana	19	Sicilia
10	Umbria	20	Sardegna

Table 3A. CAT 12 and CAT 5 classification.

CAT 12		CAT 5	
1	Agriculture, forestry and fishing	1	Agriculture
2	Manufacturing	2	Manufacturing (or Industry in the narrow sense)
3	Electricity, gas, steam and air conditioning supply	2	Manufacturing (or Industry in the narrow sense)
4	Water supply; sewerage, waste management and remediation activities	2	Manufacturing (or Industry in the narrow sense)
5	Construction	3	Construction
6	Wholesale and retail trade; repair of motor vehicles and motorcycles	4	Trade (Commerce/Wholesale & Retail)
7	Transportation and storage	5	Other service activities/Other services
8	Accommodation and food service activities	5	Other service activities/Other services
9	Information and communication	5	Other service activities/Other services
10	Financial and insurance activities	5	Other service activities/Other services
11	Real estate activities	5	Other service activities/Other services
12	Other services (professional, artistic, health, education, public administration, etc.)	5	Other service activities/Other services

Appendix 8: Shift–share instrumental variable (IV) strategy

We adopt a shift–share instrumental variable (IV) strategy (Bartik 1991; Goldsmith-Pinkham et al., 2020) to identify the causal impact of exogenous female labour supply shocks on regional and industrial changes in gender composition. The approach exploits the 2017 childcare and parental-leave reform, which acted as a national exogenous shock to female labour supply. The instrument is constructed in two steps. In the first step, we compute the national-level shock to female labor supply for each age group a in 2017, denoted by $\Delta S_{a,2017}$. In the second step, we interact these shocks with baseline female employment shares by region r , industry i (according to CAT12 classification), and age group a , measured in 2008. Let $S_{ria,2008}^F$ denote the share of employed women in age group a , industry i , and region r in 2008, obtained from microdata. The resulting shift-share instrument for each region–industry cell (r, i) is given by:

$$Z_{ri} = \sum_a S_{ria,2008}^F * \Delta S_{a,2017}$$

We then estimate the exposure of each cell (r, i) to the 2017 reform using the following regression specification:

$$\Delta f_{ri} = \alpha + \beta Z_{ri} + \gamma X_{ri} + \mu_r + \varepsilon_{ri}$$

Where Δf_{ri} denotes the change in the female employment in region r and industry i , Z_{ri} is the shift-share instrument as defined above, X_{ri} is a vector of control variables (baseline_fshare_2008 and log_F_2008),

μ_r captures region fixed effects, and ϵ_{ri} is the idiosyncratic error term. Figure 1.A in the Appendix 1 provides descriptive evidence on the heterogeneity of gender composition adjustments across regions and industries. The heatmap of changes in female employment shares between 2016 and 2020 shows marked variation: some cells display increases of more than 20 percentage points, while others record comparable declines. This motivates the use of a shift–share IV, as such differences cannot be explained by uniform national trends. Table 4.A in the Appendix displays the estimated effects on female employment share.



Figure 1A. Change in female employment share by region and industry (2016–2020).

Table 4A. Estimated effects on female employment share.

Dependent variable Δf_{ri} the change in the female employment in region r and industry i	
Variable	Coefficient
Zshiftshare2017	14.30 (0.180)
baselinefshare2008	−0.0597 (0.153)
logF_2008	0.0019 (0.601)

P value in parenthesis. Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $p < 0.1$.

The results indicate that the shift–share instrument (2017) has a positive coefficient (14.3), suggesting that a larger female labor supply shock is associated with an increase in the female share of employment across region–industry cells. However, the effect is not statistically significant ($p = 0.18$), implying that we cannot rule out the possibility that the relationship is due to chance.

The historical controls do not emerge as strong predictors either. The baseline female share in 2008 has a negative coefficient (−0.06), consistent with the idea that regions and industries with initially higher female participation experienced smaller subsequent increases. Yet, this effect is also not statistically significant ($p = 0.15$). Similarly, the initial female employment size ($\log F_{2008}$) shows a coefficient close to zero (0.0019) and is far from significant ($p = 0.60$). In sum, this specification does not provide robust evidence supporting a causal impact of the 2017 shift–share shock on changes in the female share, while confirming that initial conditions explain part of the observed variation, albeit without reaching conventional levels of statistical significance. The shift–share IV results suggest that the 2017 childcare and parental-leave reform generated a positive but statistically insignificant effect on female employment shares across region–industry cells. While this specification does not provide robust causal evidence, the direction of the estimated coefficients is consistent with our theoretical framework. Regions and industries that were more exposed to the exogenous female labour supply shock display increases in female participation, even if the effect is heterogeneous and not statistically significant. This evidence can be linked to the RV/UV measures developed in the paper. The positive, albeit weak, impact of the reform implies that incremental changes in female labour supply—particularly in industries where women are underrepresented—may enhance Related Variety (RVF) by

fostering within-industry diversification. At the same time, the broader distribution of women across industries could strengthen Unrelated Variety (UVF), thereby reinforcing portfolio effects and resilience. The absence of strong statistical significance is not inconsistent with this interpretation, as the adjustment of industrial structures and knowledge spillovers often requires longer horizons to materialize. Overall, the robustness check highlights that exogenous policy shocks can shift gender composition in ways that are directionally consistent with the mechanisms identified in our RV/UV analysis: underrepresentation creates opportunities for increasing returns through cognitive spillovers, while more balanced gender distributions across industries can amplify the benefits of diversity.

References

- Aarstad, J., Kvitastein, O. A., and Jakobsen, S. E. (2016) 'Related and Unrelated Variety as Regional Drivers of Enterprise Productivity and Innovation: A Multilevel Study', *Research Policy*, 45, 844–856. <https://doi.org/10.1016/j.respol.2015.12.013>
- Abel, J. R., Dey, I., and Gabe, T. M. (2008) 'Productivity and the Density of Human Capital', *Journal of Regional Science*, 48, 401–426. <https://doi.org/10.1111/j.1467-9787.2008.00584.x>
- Akita, T. (2003) 'Decomposing Regional Income Inequality in China and Indonesia using a Two-Stage Nested Theil Decomposition Method', *The Annals of Regional Science*, 37, 55–77. <https://doi.org/10.1007/s00168-002-0131-y>
- Alesina, A., Giuliano, P., and Nunn, N. (2013) 'On the Origins of Gender Roles: Women and the Plough', *The Quarterly Journal of Economics*, 128, 469–530. <https://doi.org/10.1093/qje/qjt005>
- Attaran, M. (1986) 'Industrial Diversity and Economic Performance in U.S. Areas', *The Annals of Regional Science*, 20, 44–54. <https://doi.org/10.1007/BF00122248>
- Balland, P.-A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P., Rigby, D. L., and Hidalgo, C. A. (2020) 'Complex Economic Activities Concentrate in Large Cities', *Nature Human Behaviour*, 4, 248–254. <https://doi.org/10.1038/s41562-019-0803-3>
- Bartik, T. J. (1991) *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W. E. Upjohn Institute for Employment Research. <https://doi.org/10.17848/9780585223940>
- Blau, F. D., and Kahn, L. M. (2017) 'The Gender Wage Gap: Extent, Trends, and Explanations', *Journal of Economic Literature*, 55, 789–865. <https://doi.org/10.1257/jel.20160995>
- Boschma, R. (2005) 'Proximity and Innovation: A Critical Assessment', *Regional Studies*, 39, 61–74. <https://doi.org/10.1080/0034340052000320887>
- Boschma, R., and Iammarino, S. (2009) 'Related Variety, Trade Linkages, and Regional Growth in Italy', *Economic Geography*, 85, 289–311. <https://doi.org/10.1111/j.1944-8287.2009.01034.x>
- Boschma, R., and Frenken, K. (2011) 'The Emerging Empirics of Evolutionary Economic Geography', *Journal of Economic Geography*, 11, 295–307. <https://doi.org/10.1093/jeg/lbq053>
- Boschma, R., and Gianelle, C. (2014) Regional branching and smart specialisation policy. JRC Scientific and Policy Reports, JRC88242. European Commission, Joint Research Centre. <https://doi.org/10.2791/93300>
- Bowers, C. A., Pharmer, J. A., and Salas, E. (2000) 'When Member Homogeneity is Needed in Work Teams: A Meta-Analysis', *Small Group Research*, 31, 305–327. <https://doi.org/10.1177/104649640003100303>
- Castaldi, C., Frenken, K., and Los, B. (2015) 'Related Variety, Unrelated Variety and Technological Breakthroughs: An Analysis of US State-Level Patenting', *Regional Studies*, 49, 767–781. <https://doi.org/10.1080/00343404.2014.940305>
- Das, S., and Kotikula, A. (2018) *Gender-Based Employment Segregation: Understanding Causes and Policy Interventions*. Washington, DC: World Bank. <https://doi.org/10.1596/30695>
- Duncan, O. D., and Duncan, B. (1955) 'A Methodological Analysis of Segregation Indexes', *American Sociological Review*, 20, 210–217. <https://doi.org/10.2307/2088328>
- Felin, T., Foss, N. J., Heimeriks, K. H., and Madsen, T. L. (2012) 'Microfoundations of Routines and Capabilities: Individuals, Processes, and Structure', *Journal of Management Studies*, 49, 1351–1374. <https://doi.org/10.1111/j.1467-6486.2012.01052.x>
- Frenken, K., Van Oort, F., and Verburg, T. (2007) 'Related Variety, Unrelated Variety and Regional Economic Growth', *Regional Studies*, 41, 685–697. <https://doi.org/10.1080/00343400601120296>
- Fritsch, M., and Kublina, S. (2018) 'Related Variety, Unrelated Variety, and Regional Growth: The Role of Absorptive Capacity and Entrepreneurship', *Regional Studies*, 52, 1360–1371. <https://doi.org/10.1080/00343404.2017.1399630>

- Giannantoni, C., and Rodríguez-Pose, A. (2025) 'Regional Government Institutions and the Capacity for Women to Reconcile Career and Motherhood', *Journal of Economic Geography*, 25, 311–328. <https://doi.org/10.1093/jeg/lbad029>
- Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020) 'Bartik Instruments: What, When, Why, and How', *American Economic Review*, 110, 2586–2624. <https://doi.org/10.1257/aer.20181047>
- Grant, R. M. (1996) 'Toward a Knowledge-Based Theory of the Firm', *Strategic Management Journal*, 17, 109–122. <https://doi.org/10.1002/smj.4250171110>
- Griffin, D., Li, K., and Xu, T. (2021) 'Board Gender Diversity and Corporate Innovation: International Evidence', *Journal of Financial and Quantitative Analysis*, 56, 123–154. <https://doi.org/10.1017/S0022109020000119>
- Hackbart, M. M., and Anderson, D. A. (1975) 'On Measuring Economic Diversification', *Land Economics*, 51, 374–378. <https://doi.org/10.2307/3145727>
- Hakim, C. (2000) *Work-Lifestyle Choices in the 21st Century: Preference Theory*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199242109.001.0001>
- Hall, J. V., and Krueger, A. B. (2018) 'An Analysis of the Labor Market for Uber's Driver-Partners in the United States', *ILR Review*, 71, 705–732. <https://doi.org/10.1177/0019793917717222>
- Hernández-Lara, A. B., González-Bustos, J. P., and Alarcón-Alarcón, A. (2021) 'Social Sustainability on Corporate Boards: The Effects of Female Family Members on R&D', *Sustainability*, 13, 1913. <https://doi.org/10.3390/su13041913>
- Hidalgo, C. A., Klinger, B., Barabási, A.-L., and Hausmann, R. (2007) 'The Product Space Conditions the Development of Nations', *Science*, 317, 482–487. <https://doi.org/10.1126/science.1144581>
- Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019) 'The Allocation of Talent and U.S. Economic Growth', *Econometrica*, 87, 1439–1474. <https://doi.org/10.3982/ECTA11427>
- Jackson, S. E., Brett, J. F., Sessa, V. I., Cooper, D. M., Julin, J. A., and Peyronnin, K. (1995) 'Some Differences Make a Difference: Individual Dissimilarity and Group Heterogeneity as Correlates of Recruitment, Promotions, and Turnover', *Journal of Applied Psychology*, 80, 675–689. <https://doi.org/10.1037/0021-9010.80.5.675>
- Jacobs, J. (1961) *The Death and Life of Great American Cities*. New York: Random House.
- Jacobs, J. (1969) *The Economy of Cities*. New York: Random House.
- Karmel, T., and MacLachlan, M. (1988) 'Occupational Sex Segregation—Increasing or Decreasing?', *The Economic Record*, 64, 187–195. <https://doi.org/10.1111/j.1475-4932.1988.tb02010.x>
- Karakowsky, L., and Siegel, J. P. (1999) 'The Effects of Proportional Representation and Gender Orientation on Emergent Leadership Behaviors in Mixed-Gender Work Groups', *Journal of Applied Psychology*, 84, 620–631. <https://doi.org/10.1037/0021-9010.84.4.620>
- Martini, B. (2020) 'Resilience and Economic Structure: Are They Related?', *Structural Change and Economic Dynamics*, 54, 62–91. <https://doi.org/10.1016/j.strueco.2020.03.001>
- Neffke, F., and Henning, M. (2013) 'Skill Relatedness and Firm Diversification', *Strategic Management Journal*, 34, 297–316. <https://doi.org/10.1002/smj.2014>
- Nelson, R. R., and Winter, S. G. (1982) *An Evolutionary Theory of Economic Change*. Cambridge, MA: Belknap Press of Harvard University Press.
- Ngai, L. R., and Petrongolo, B. (2017) 'Gender Gaps and the Rise of the Service Economy', *American Economic Journal: Macroeconomics*, 9, 1–44. <https://doi.org/10.1257/mac.20150253>
- Nielsen, M. W., Alegria, S., Börjeson, L., Etkowitz, H., Falk-Krzesinski, H. J., Joshi, A., Leahey, E., Smith-Doerr, L., Woolley, A. W., and Schiebinger, L. (2017) 'Gender Diversity Leads to Better Science', *Proceedings of the National Academy of Sciences*, 114, 1740–1742. <https://doi.org/10.1073/pnas.1700616114>
- Nooteboom, B. (2000) 'Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance', *Journal of Management and Governance*, 4, 69–92. <https://doi.org/10.1023/A:1009941416749>
- Olivetti, C., and Petrongolo, B. (2016) 'The Evolution of Gender Gaps in Industrialized Countries', *Annual Review of Economics*, 8, 405–434. <https://doi.org/10.1146/annurev-economics-080614-115329>
- O'Reilly, C. A., Williams, K. Y., and Barsade, S. (1997) 'Group Demography and Innovation: Does Diversity Help?', In E., Mannix, and M. Neale (eds.) *Research in the Sociology of Organizations*, Vol. 19, pp. 183–207. Greenwich, CT: JAI Press.
- Østergaard, C. R., Timmermans, B., and Kristinsson, K. (2011) 'Does a Different View Create Something New? The Effect of Employee Diversity on Innovation. *Research Policy*, 40, 500–509. <https://doi.org/10.1016/j.respol.2010.11.004>
- Petrongolo, B., and Ronchi, M. (2020) 'Gender Gaps and the Structure of Local Labor Markets', *Labour Economics*, 64, 101819. <https://doi.org/10.1016/j.labeco.2020.101819>

- Rodríguez-Pose, A. (2018) 'The Revenge of the Places that Don't Matter (and What To Do About It)', *Cambridge Journal of Regions, Economy and Society*, 11, 189–209. <https://doi.org/10.1093/cjres/rsx024>
- Rodríguez-Pose, A., and Ketterer, T. (2020) 'Institutional Change and the Development of Lagging Regions in Europe', *Regional Studies*, 54, 974–986. <https://doi.org/10.1080/00343404.2019.1608356>
- Rodríguez-Pose, A., and Ganau, R. (2022) 'Institutions and the Productivity Challenge for European Regions', *Journal of Economic Geography*, 22, 1–25. <https://doi.org/10.1093/jeg/lbab028>
- Sastre, R. (2015) 'Gender Diversity and Innovation: The Role of Women in R&D Teams', *International Journal of Entrepreneurship and Small Business*, 24, 142–162. <https://doi.org/10.1504/IJESB.2015.066154>
- Shannon, C. E. (1948) 'A Mathematical Theory of Communication', *The Bell System Technical Journal*, 27, 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Simon, H. A. (1991) 'Organizations and Markets', *Journal of Economic Perspectives*, 5, 25–44. <https://doi.org/10.1257/jep.5.2.25>
- Straathof, S. M. (2007) 'Shannon's Entropy as an Index of Product Variety', *Economics Letters*, 94, 297–303. <https://doi.org/10.1016/j.econlet.2006.09.022>
- Ter Wal, A. L. J., and Boschma, R. A. (2011) 'Co-Evolution of Firms, Industries and Networks in Space', *Regional Studies*, 45, 919–933. <https://doi.org/10.1080/00343400802662658>
- Theil, H. (1965) *Economic Forecasts and Policy*, 2nd edn. Amsterdam: North-Holland.
- Theil, H. (1967) *Economics and Information Theory*. Amsterdam: North-Holland.
- Torchia, M., Calabrò, A., Gabaldon, P., and Kanadli, S. B. (2018) 'Women Directors' Contributions to Organizational Innovation: A Behavioral Approach', *Scandinavian Journal of Management*, 34, 215–224. <https://doi.org/10.1016/j.scaman.2018.02.001>
- Valantine, H. A., and Collins, F. S. (2015) 'National Institutes of Health Addresses the Science of Diversity', *Proceedings of the National Academy of Sciences*, 112, 12240–12242. <https://doi.org/10.1073/pnas.1515612112>
- Williams, K. Y., and O'Reilly, C. A. (1998) 'Demography and Diversity in Organizations: A Review of 40 Years of Research', *Research in Organizational Behavior*, 20, 77–140.
- Wood, W., Polek, D. B., and Aiken, L. S. (1985) 'Sex Differences in Group Task Performance', *Journal of Personality and Social Psychology*, 48, 63–71. <https://doi.org/10.1037/0022-3514.48.1.63>
- Wu, D., Yuan, L., Li, R., and Li, J. (2018) 'Decomposing Inequality in Research Funding by University–Institute Sub-Group', *Journal of Informetrics*, 12, 1312–1326. <https://doi.org/10.1016/j.joi.2018.09.003>
- Xie, L., Wang, S., and Zhang, Y. (2020) 'Gender Diversity in R&D Teams and Innovation Efficiency', *Research Policy*, 49, 103921. <https://doi.org/10.1016/j.respol.2020.103921>