The Pink-Collar Trap: Occupational Segregation and Gendered Vulnerability to AI

Thais Palanca[1]

This research was supported by NOVA School of Business and Economics and by the Fundação para a Ciência e a Tecnologia (FCT).[2]

Abstract

The occupations that brought women into the twentieth-century labor force (clerks, cashiers, receptionists, sales assistants) are now the most vulnerable to artificial intelligence. In Brazil, half of young women entering the labor market work in these AI-substitutable roles, compared to 29% of young men. A general equilibrium model with occupational sorting, switching costs, and educational barriers predicts that full AI adoption widens the within-cohort (entrants) gender earnings gap by 2.43 percentage points, driven overwhelmingly by wages (83%): young women absorb declining earnings in disrupted occupations while young men, concentrated in less-exposed sectors, are largely insulated. But here is the puzzle: the aggregate gender gap barely moves. The reason is that incumbent women, who hold professional positions where AI augments rather than replaces, benefit from the same technology that hurts entrants. This compositional masking effect means that any analysis pooling age groups will conclude, incorrectly, that AI is gender-neutral. It is not. Retraining 35% of non-degree women in vulnerable occupations eliminates the gap widening entirely.

Keywords: Artificial Intelligence, Gender Inequality, Occupational Segregation, Labor Force Participation, Entry Dynamics
JEL Codes: J16, J22, J24, J31, O33

1 Introduction

The rapid advancement of artificial intelligence, particularly generative AI systems, raises important questions about the future distribution of labor market opportunities. While previous waves of automation primarily affected manual and routine tasks, AI increasingly performs cognitive tasks traditionally considered safe from technological displacement (Acemoglu and Restrepo, 2022; Autor, 2022). A growing literature examines which occupations are most exposed to AI, but less attention has been paid to *who* works in these occupations and how AI-related risks are distributed across demographic groups.

This paper documents that AI vulnerability in Brazil is strongly concentrated among young women. Using PNADc 2023 microdata merged with occupation-level AI exposure and com-

[1]PhD Candidate, NOVA School of Business and Economics. Email: thais.silva@novasbe.pt

[2]I gratefully acknowledge financial support from NOVA School of Business and Economics and from the Fundação para a Ciência e a Tecnologia (FCT), Portugal.

plementarity measures from Cazzaniga et al. (2024), I find that 50.5% of women aged 18–25 work in High Exposure, Low Complementarity (HELC) occupations—where AI substitutes rather than augments human labor—compared to 28.8% of young men. Four pink-collar occupations (shop sales assistants, general office clerks, cashiers, and receptionists) account for 86% of this gap, and the concentration is sharpest at labor market entry, consistent with gender-differentiated occupational sorting at the point of career formation. Exploiting the PNADc's rotating panel to observe quarter-to-quarter transitions for 376,293 individuals, I find that household structure—marriage and children—explains more than two-thirds of the gender gap in labor force exit. Marriage generates a 36 percentage point reversal: it pushes men into the labor force but pulls women out, consistent with household specialization rather than spousal insurance. These patterns are structurally stable across 2015–2023, confirming that the 2023 calibration is representative.

I develop a static general equilibrium model with four demographic groups, three sectors, occupational switching costs, and educational barriers, building on Ngai and Petrongolo (2017). Workers sort across sectors based on comparative advantage (Roy, 1951; Heckman and Sedlacek, 1985), with outside options that vary by household type, calibrated from the participation analysis in which marriage and children were identified as the dominant determinants of gendered labor force exit. The model predicts that full AI adoption widens the within-cohort gender earnings gap by 2.4 percentage points, driven primarily by the wage channel (83%): young women absorb declining HELC wages while young men, concentrated in less-exposed sectors, are largely insulated. A compositional masking effect emerges from the model: incumbent women, who hold 21.5% of high-complementarity professional positions, benefit from AI, offsetting the losses of young entrants and leaving the aggregate gender gap nearly unchanged. This means that policies designed around aggregate gender statistics will miss the cohort that needs intervention most. Counterfactual experiments targeting young entrants specifically show that retraining 35% of non-degree women in vulnerable occupations eliminates the gap widening, while occupational desegregation of four specific occupations captures most of this effect.

The paper makes four contributions. First, to the AI and labor markets literature, I document that aggregate gender analyses of AI's impact are misleading: young women entrants lose while incumbent women gain, but pooling age groups masks this heterogeneity entirely. Second, to the gender and technology literature, I show that the same pink-collar occupations that enabled women's economic advancement in the twentieth century (clerical work, retail, administrative support) now constitute the principal source of gendered AI vulnerability. Third, to the household labor supply literature, I demonstrate that marriage and children explain more than two-thirds of the gender exit gap, but that the general equilibrium model identifies wages, not exit, as the dominant channel through which AI widens the earnings gap—a finding that reverses the conventional prior. Fourth, to the structural modeling literature, I develop a tractable GE framework that generates occupational sorting, participation, and the gender exit ratio as endogenous outcomes from estimated primitives, and show that the educational barrier—not preferences or norms—is what makes the pink-collar trap binding.

The remainder of the paper proceeds as follows. Section 2 reviews related literature in detail. Section 3 describes the data and methodology. Section 4 presents empirical findings on the demographic distribution of AI vulnerability. Section 5 develops the theoretical framework. Section 6 presents calibration and quantitative results. Section 7 concludes.

2 Related Literature

This paper contributes to several interconnected literatures on technology, labor markets, and gender inequality. I discuss each in turn and highlight how this paper extends existing research.

2.1 AI and Labor Market Disruption

The literature on technology and labor markets has evolved substantially over the past two decades. The canonical framework of Autor et al. (2003) emphasized routine-biased technological change, wherein computers substitute for workers performing routine cognitive and manual tasks while complementing those performing non-routine abstract tasks. This framework successfully explained the job polarization observed in developed economies during the 1990s and 2000s (Goos and Manning, 2007; Autor, 2015), as employment shifted away from middle-skill routine occupations toward both high-skill cognitive work and low-skill manual services.

However, the emergence of artificial intelligence challenges this paradigm. Brynjolfsson and McAfee (2014) argue that AI represents a "second machine age" capable of performing cognitive tasks previously considered uniquely human. Unlike earlier automation technologies that required explicit programming, machine learning systems can recognize patterns, process natural language, and make decisions under uncertainty—capabilities that were thought to be safely non-routine (Autor, 2022). This has led to concerns that AI may affect a broader swath of the occupational distribution than previous technological waves.

A growing literature develops measures of occupational exposure to AI. Felten et al. (2021) construct an AI Occupational Exposure index by linking specific AI capabilities (such as image recognition, language modeling, and speech recognition) to occupational task requirements from O*NET. Their approach reveals that AI exposure is highest in professional and technical occupations, contrasting with routine-biased technological change which primarily affected middle-skill workers. Webb (2020) uses natural language processing to match AI patents to occupational descriptions, finding that high-skill occupations face greater exposure than under previous automation waves. Eloundou et al. (2023) specifically examine exposure to large language models, estimating that approximately 80% of the U.S. workforce could have at least 10% of their work tasks affected by GPT-4 and related systems.

A crucial insight from recent work is that AI exposure alone does not determine labor market impacts. Cazzaniga et al. (2024) introduce the concept of *complementarity*, which captures whether AI augments worker productivity or substitutes for human labor in a given occupation. The complementarity index $\theta \in [0, 1]$ is constructed by analyzing task content: occupations score high on complementarity when their core tasks involve judgment, creativity, interpersonal interaction, or physical presence that AI cannot easily replicate, even if AI can assist with peripheral tasks. Occupations score low on complementarity when AI can perform their core tasks—such as data entry, routine document processing, or standardized customer service—more cheaply and at scale.

This framework resolves an apparent paradox in the exposure-based literature: some high-exposure occupations may benefit from AI while others face displacement. For example, software developers and financial analysts both have high AI exposure, but developers score high on complementarity (AI assists but does not replace creative problem-solving) while certain

analyst roles score low (AI can automate routine data processing and report generation). The key insight is that *the combination of exposure and complementarity* determines labor market outcomes, not exposure alone.

This allows us to classify occupations into three categories that I adopt in this paper following Cazzaniga et al. (2024): Low Exposure (LE), where AI capabilities have limited overlap with job tasks; High Exposure with High Complementarity (HEHC), where AI augments worker productivity; and High Exposure with Low Complementarity (HELC), where AI substitutes for human labor. Their analysis of 30 countries reveals that approximately 40% of global employment is exposed to AI, with exposure higher in advanced economies. Crucially, they find substantial heterogeneity within the high-exposure category: about half of exposed workers are in HEHC occupations (likely to benefit) while half are in HELC occupations (likely to face displacement).

The complementarity concept is central to this paper for two reasons. First, it reveals that gender differences in AI vulnerability are larger than exposure-based measures suggest. Women have higher average AI exposure than men, but they also have *lower* average complementarity—meaning their exposure is concentrated in occupations where AI substitutes rather than augments. Second, the complementarity framework provides the theoretical foundation for my model's sector-specific productivity effects: HEHC workers experience wage gains as AI augments their productivity, while HELC workers face wage losses and potential displacement.

While this literature has made substantial progress in measuring which occupations are exposed to AI, far less attention has been paid to *who* works in these occupations. Gmyrek et al. (2023) provide an important insight, showing that clerical work is the only broad occupational category with high automation potential and noting that the effects are strongly gendered given women's overrepresentation in clerical roles, particularly in upper-middle-income countries. Cazzaniga et al. (2025) extend this analysis across six economies, finding that women in high-substitution occupations at the bottom of the wage distribution transition to inactivity more frequently. However, neither study examines the within-country demographic structure, specifically, how age $\times$ gender interactions generate heterogeneous AI vulnerability that is masked in aggregate comparisons. This paper addresses that gap.

2.2 Gender, Technology, and Occupational Segregation

Historically, technological change has often reduced gender inequality in labor markets. Goldin (2006) describes the "quiet revolution" in women's labor force participation facilitated by technological change in home production and contraception. Black and Spitz-Oener (2008) document how technological advances in the 20th century narrowed gender gaps by reducing the premium on physical strength, enabling women's entry into clerical and service occupations. The typewriter, telephone, and personal computer created new occupational categories that women entered in large numbers.

The routine-biased technological change era produced mixed effects. Cortés and Pan (2020) show that the decline of routine employment disproportionally affected men in routine manual occupations, while women in non-routine service and cognitive occupations were relatively protected. However, Beaudry et al. (2016) document a "great reversal" in the demand for cognitive skills after 2000, with negative implications for women's relative wages.

Several studies suggest that AI may reverse historical patterns. Brussevich et al. (2018) find that women face higher average AI exposure across 30 countries because they are concentrated in cognitive, non-routine occupations now threatened by AI capabilities. Del Rio-Chanona et al. (2021) similarly find that female-dominated occupations have higher LLM exposure, with administrative support, education, and healthcare particularly affected.

The mechanism underlying these patterns is occupational segregation. Despite decades of progress, labor markets remain highly segregated by gender (Blau and Kahn, 2017; Cortés and Pan, 2018). Goldin (2014) argues that the remaining gender gap reflects occupational and industry segregation more than within-occupation discrimination—a finding with direct implications for how AI differentially affects men and women. The historical origins of this segregation are well documented. Goldin (1984) describes how clerical occupations absorbed large numbers of female labor force entrants in the early twentieth century, transforming from male-dominated to female-dominated within a few decades. Howe (1977) coined the term *pink-collar* to describe these female-dominated occupations—secretaries, receptionists, retail clerks, cashiers—that offered women entry into the paid labor force but at lower wages and with limited advancement. The pink-collar category has since expanded to encompass a broader set of service and administrative roles characterized by routine cognitive tasks, interpersonal interaction, and female overrepresentation (Blau and Kahn, 2017).

Importantly, the previous wave of automation partially *unwound* this segregation. Cortés et al. (2024) show that routine-biased technological change driven by personal computers reduced occupational segregation in the United States, as women moved out of routine clerical and manufacturing positions and into cognitive-analytical roles. AI threatens to reverse this progress: whereas computerization displaced routine tasks while complementing abstract cognitive work, generative AI directly targets the cognitive tasks in which women concentrated after the computer revolution. The pink-collar occupations that absorbed women leaving routine manufacturing now face substitution themselves.

This pink-collar concentration is precisely what generates gendered AI vulnerability. The same occupational pathways that enabled women's economic advancement in the 20th century now exhibit the lowest human-AI complementarity. In Brazil, I document that four pink-collar occupations—shop sales assistants, general office clerks, cashiers, and receptionists—account for 86% of the gender gap in AI vulnerability among young workers. These occupations share a common feature: they involve routine cognitive tasks (processing transactions, managing information, answering queries) that large language models and AI assistants can increasingly perform, but they lack the complementarities with AI that characterize professional and technical work.

The closest existing work is Cazzaniga et al. (2025), who examine gendered AI exposure across six economies including Brazil. They document that women in high-substitution occupations at the bottom of the wage distribution transition to inactivity more frequently. I extend their analysis in three directions: I decompose the demographic incidence by age $\times$ gender (revealing the aggregation paradox they do not examine), I develop a structural GE model that endogenizes both sorting and exit, and I provide policy counterfactuals quantifying the effect of retraining and occupational desegregation.

This paper extends the gender and technology literature in three ways. First, I show that AI exposure alone understates gender differences — what matters is the combination of high

exposure and low complementarity that characterizes pink-collar work. Second, rather than documenting exposure descriptively, I develop a general equilibrium model in which occupational sorting emerges endogenously from comparative advantage, allowing evaluation of counterfactuals that alter the sorting itself. Third, I document a compositional masking effect that existing studies miss: young women entrants lose from AI while incumbent women in professional occupations benefit, leaving the aggregate gender gap nearly unchanged. Analyses that pool age groups will conclude, incorrectly, that AI is gender-neutral.

The focus on entrants is motivated by scarring literature. Oreopoulos et al. (2012) and Kahn (2010) document that initial occupational conditions have persistent career effects, what suggests that AI's disruption of traditional entry pathways may generate inequality that extends beyond the immediate shock. In Brazil I show that four pink-collar occupations account for 86% of the gender gap AI vulnerability among young workers, providing a concrete target for policy interventions.

2.3 Household Structure and the Extensive Margin

A substantial literature documents that married and single individuals sort into different occupations, with patterns that differ by gender. Korenman and Neumark (1991) document a "marriage premium" for men, potentially reflecting selection, specialization, or employer discrimination. For women, the "motherhood penalty" (Budig and England, 2001; Kleven et al., 2019) and household specialization push married women toward occupations that offer temporal flexibility at the cost of lower wages (Goldin, 2014). This sorting mechanism implies that household structure may shape not only where workers are located in the occupational distribution but also how they respond when AI disrupts their sector. No study to my knowledge has examined how household structure determines the *exit response* to AI-driven occupational disruption.

The canonical labor supply model with selection (Heckman, 1979) posits that individuals participate in the labor market when their market wage exceeds their reservation wage. This implies that observed wages are conditional on participation, so changes in the wage gap may reflect changes in selection rather than in the wage structure (Mulligan and Rubinstein, 2008). Olivetti and Petrongolo (2008) show that this selection effect can reverse cross-country rankings of gender inequality, motivating the use of *earnings* gaps which incorporate both wages and participation rather than wage gaps alone. Surprisingly little work examines how technological change affects the extensive margin of labor supply; Autor (2019) documents declining male participation linked to automation, but the gendered implications remain underexplored.

This paper makes the wage-earnings distinction central. Using participation probits estimated from PNADc microdata, I show that marriage and children explain more than two-thirds of the gender gap in labor force exit, with marriage generating a 36 percentage point reversal between men and women — consistent with household specialization rather than spousal insurance. However, the general equilibrium model reveals that the wage channel, not exit, is the primary mechanism through which AI widens the gender earnings gap, accounting for 83% of the within-cohort effect. Most observed gender differences in inactivity reflect baseline household specialization rather than marginal responses to AI-induced wage declines. This redirects policy attention: interventions that change where women work, such as retraining and occupational access, target the dominant channel, while policies such as childcare subsidies target the

secondary but non-negligible exit channel.

2.4 General Equilibrium Models of Gender and Structural Change

A growing literature develops general equilibrium models to study how structural change interacts with gender inequality. This paper contributes to this tradition. The most closely related framework is Ngai and Petrongolo (2017), who model gendered sectoral sorting in a two-sector economy where structural change (the rise of services) narrows gender gaps by expanding the sector in which women have comparative advantage. I build on their structure but introduce three sectors classified by AI complementarity, replace secular structural change with an AI productivity shock, and add four demographic groups that reveal heterogeneous effects across age and gender. Rendall (2018) and Cerina et al. (2021) extend this framework to study how technological change interacts with gendered occupational sorting over the long run, documenting that the rise of the service sector explaining a substantial share of the convergence in gender gaps across developed economies. My analysis complements theirs by asking whether AI could reverse this convergence for a specific group, as it has the potential of disrupting the services sector rather than expanding it.

Hsieh et al. (2019) model barriers by race an gender to occupational entry generate misallocation of talent and reduce aggregate output. I add an educational barrier for HEHC entry, following their logic: workers who could have been productive in high complementarity occupations are locked out by insufficient credentials, and by tackling this barrier, I found the most effective intervention in the model.

The model also adapts the quantitative spatial and trade literature, particularly Caliendo et al. (2019), who develop a GE framework with worker mobility across sectors and regions subject to switching costs. I adopt a similar structure, with heterogeneous workers choosing among discrete options with idiosyncratic taste shocks and mobility frictions, but apply it to gender and AI rather than trade and geography. This approach builds on Traiberman (2019), who estimates a model of occupational choice with switching costs calibrated from observed transition rates in the context of Danish workers responding to import competition. I adopt his calibration strategy for group-specific switching costs using PNADc quarterly transition data. Relative to this literature, the main contribution is the aggregation finding: by modeling four demographic groups rather than two genders, the model reveals that AI's gender impact is concentrated among labor market entrants and is obscured in aggregate comparisons — a result that two-gender or representative-agent models cannot produce.

3 Data and Methodology

3.1 Data Sources

The analysis combines two data sources. The primary source is the Pesquisa Nacional por Amostra de Domicílios Contínua (PNADc) 2023, Brazil's main continuous household survey conducted quarterly by IBGE. The PNADc employs a rotating panel design in which each household is interviewed for five consecutive quarters, enabling observation of individual-level occupational transitions. I pool all four quarters of 2023, restricting attention to employed indi-

viduals aged 18–65 with valid occupation codes, yielding 759,786 person-quarter observations representing approximately 367 million weighted person-quarters[3].

The rotating panel structure provides a methodological advantage over cross-sectional data: rather than inferring transition patterns from age-group comparisons, I directly observe quarter-to-quarter occupational changes for the same individuals. I construct individual identifiers from the combination of state, primary sampling unit, household number, person number within household, sex, and birth year. Restricting to consecutive-quarter transitions (quarter difference of one) yields 376,293 valid individual transitions.

I merge the PNADc microdata with occupation-level measures of AI exposure and complementarity from Cazzaniga et al. (2024). The AI Occupational Exposure Index (AIOE), originally developed by Felten et al. (2021), measures the degree to which AI capabilities overlap with occupational task requirements based on O*NET task content. The complementarity index captures whether AI is likely to augment worker productivity or substitute for human labor.

3.2 Variable Construction

AI Risk Categories. Following the framework developed by Pizzinelli et al. (2023) and applied to Brazil by Cazzaniga et al. (2024b), I classify occupations into three mutually exclusive categories based on median splits of exposure and complementarity among high-exposure occupations:

- *Low Exposure (LE)*: AI exposure below the sample median. These occupations face limited disruption regardless of complementarity.
- *High Exposure, High Complementarity (HEHC)*: Exposure at or above the median with complementarity at or above the conditional median. AI is likely to augment productivity in these occupations.
- *High Exposure, Low Complementarity (HELC)*: Exposure at or above the median with complementarity below the conditional median. These are high-risk occupations where AI is likely to substitute for human labor.

Cazzaniga et al. (2024b) apply this classification to PNADc data and document that college-educated workers in Brazil frequently transition from HELC to HEHC over their careers, while non-college workers face higher risks of downward mobility. I extend their analysis by decomposing the demographic incidence by age $\times$ gender and by developing a structural model to quantify the equilibrium implications of these patterns.

Demographic Groups. I define four groups based on gender and career stage:

- *Women Entrants (WE)*: Women aged 18–25

[3]Individuals may appear in multiple quarters due to the rotating panel design. For descriptive statistics, pooling reduces sampling variance. For regressions, standard errors are clustered at the state level (27 units). Appendix E.7 confirms that single-quarter estimates are nearly identical (maximum difference $\approx$ 2 pp), indicating that pooling does not bias the results.

- *Men Entrants (ME)*: Men aged 18–25
- *Women Incumbents (WI)*: Women aged 26–65
- *Men Incumbents (MI)*: Men aged 26–65

Income. Monthly labor income is reported directly in the PNADc. I use nominal values from 2023. All regressions employ survey calibration weights (variable V1028) and cluster standard errors at the state level (27 units).

4 Empirical Results

This section documents the distribution of AI risk across demographic groups and identifies the mechanisms underlying gendered AI vulnerability.

4.1 Distribution of AI Risk

Figure 1 presents the share of each demographic group across AI risk categories. Women entrants have the highest concentration in HELC occupations at 50.5%, compared to 28.8% for men entrants, 30.4% for women incumbents, and 21.6% for men incumbents. The 21.7 percentage point gap between young women and young men constitutes the central empirical finding of this paper: more than half of young women entering the Brazilian labor market work in occupations classified as highly exposed to AI with low complementarity.

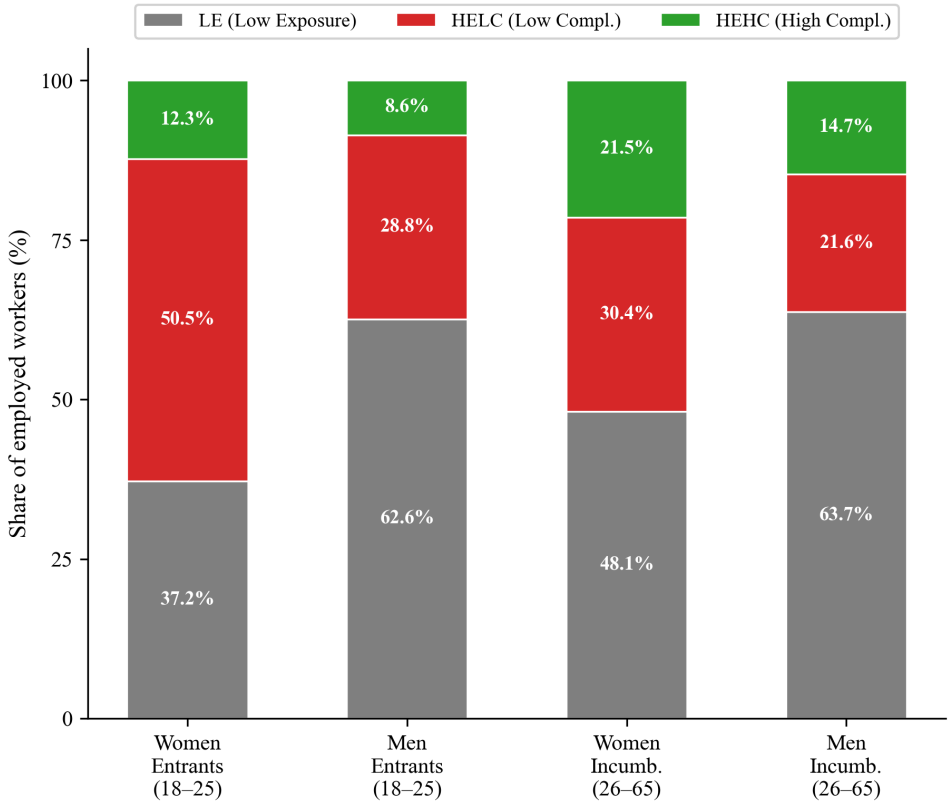


Figure 1: Distribution by AI Risk Category

Notes: Sample includes employed workers aged 18–65 from PNADc 2023 (all four quarters). HELC = High Exposure, Low Complementarity (highest substitution risk). HEHC = High Exposure, High Complementarity (AI-augmented). LE = Low Exposure.

Table 1 provides summary statistics that reveal a paradox requiring explanation. Young women have slightly more years of education than young men (12.5 vs. 11.6 years), yet they exhibit the lowest mean complementarity ($\theta = 0.528$) and highest HELC concentration among all groups. Education, which typically protects workers from technological displacement, does not insulate young women from AI risk. This pattern reflects occupational sorting: young women's educational attainment channels them toward service and clerical occupations that are uniformly classified as high-risk, while young men with comparable education enter craft, technical, and operative occupations with lower exposure.

Table 1: Summary Statistics by Demographic Group

	Entrants (18–25)		Incumbents (26–65)	
	Women	Men	Women	Men
N (millions)	25.5	33.3	135.8	172.3
Mean Age	22.0	22.0	41.9	42.3
Mean Years of Education	12.5	11.6	12.0	10.7
% in HELC	**50.5**	28.8	30.4	21.6
Mean Complementarity (θ)	0.528	0.543	0.540	0.562
Mean Monthly Income (R$)	1,581	1,798	2,676	3,370
Labor Force Participation (%)	45.7	61.5	53.5	75.5

Notes: PNADc 2023 (Q1–Q4). Population N is weighted using V1028. LFP calculated as share of working-age population employed or actively seeking work.

The income patterns in Table 1 also merit attention. Young women earn 12% less than young men (R$1,581 vs. R$1,798) and 53% less than incumbent men (R$3,370). These baseline income gaps will be amplified under AI adoption as groups experience differential productivity shocks based on their occupational distributions. Labor force participation exhibits a similar hierarchy: young women participate at the lowest rate (45.7%), while incumbent men participate at the highest (75.5%). This 30 percentage point gap in baseline participation implies that AI-induced changes in the extensive margin can generate large effects on the earnings distribution.

4.2 The Occupational Channel

What drives young women's concentration in HELC occupations? Table 2 presents employment distributions by ISCO major group alongside AI risk characteristics.

Table 2: Occupational Distribution and AI Risk by Demographic Group

Occupation (ISCO Major Group)	Employment Share (%)				% HELC
	Women Entrants	Men Entrants	Women Incumb.	Men Incumb.	
Services and Sales	**38.8**	18.0	28.0	16.8	varies
Clerical Support	**21.5**	11.0	10.4	4.8	**100.0**
Elementary	10.1	22.2	19.6	12.1	varies
Technicians	10.5	9.2	8.9	8.4	varies
Professionals	9.7	5.2	19.0	10.1	varies
Craft and Trades	3.9	16.6	5.2	20.8	low
Machine Operators	2.8	11.3	2.8	15.1	low
Agriculture	1.4	5.1	2.6	7.8	low
Managers	1.2	1.3	3.6	4.1	low

Notes: Employment shares sum to 100% within each group. % HELC indicates the share of workers in each major group classified as High Exposure, Low Complementarity at the 4-digit occupation level.

The gender gap in AI vulnerability is concentrated in two occupation groups. First, *clerical support*: 21.5% of young women work in clerical occupations, compared to 11.0% of young men and just 4.8% of incumbent men. Clerical support is uniformly classified as HELC—100% of clerical workers fall into the high-risk category—with the lowest mean complementarity ($\theta = 0.465$) among all major groups. Second, *services and sales*: 38.8% of young women work in service and sales occupations, the largest employment category for this group. While services and sales is not uniformly HELC, specific sub-occupations within it—shop sales assistants, cashiers, and food service counter attendants—are classified as high-risk.

While Table 2 identifies the broad occupational categories, Table 3 pinpoints the specific 4-digit occupations driving the gap.

Table 3: Top Occupations Contributing to the HELC Gap (WE − ME)

ISCO Code	Occupation	% WE	% ME	HELC Rate	Gap (p.p.)
5223	Shop sales assistants	11.3	5.3	100%	+6.1
4110	General office clerks	9.7	4.8	100%	+4.9
5230	Cashiers and ticket clerks	5.1	1.1	100%	+4.0
4226	Receptionists	4.5	0.7	100%	+3.7
Subtotal (top 4)					**+18.7**
Total gap					+21.7
Share explained by top 4					*86%*

Notes: Contribution calculated as $(\pi_{WE,o} - \pi_{ME,o}) \times H_o$, where π is employment share and H_o is the HELC rate in occupation o. All four occupations have 100% HELC classification.

Four occupations account for 86% of the 21.7 percentage point gap. These occupations share a common profile: they are front-facing, customer-service roles involving routine cognitive tasks—processing transactions, managing correspondence, greeting visitors, organizing files. These are precisely the tasks where AI systems have demonstrated strong performance (Eloundou et al., 2023). Shop sales assistants alone contribute 6.1 percentage points, reflecting an employment share of 11.3% among young women versus 5.3% among young men. General office clerks contribute 4.9 percentage points, cashiers 4.0, and receptionists 3.7. Young men, by contrast, are distributed more heavily across craft and trades (16.6%), machine operators (11.3%), and elementary occupations (22.2%)—occupation groups with low AI exposure.

The historical origins of this pattern are well documented. As Goldin (2014) describes, clerical and service occupations absorbed large numbers of female labor force entrants throughout the twentieth century. The typewriter, telephone, and filing systems created demand for workers with specific skills—literacy, attention to detail, communication abilities—that women possessed but that had limited labor market value before these technologies emerged. These findings suggest that the same occupational segregation that enabled women's economic advancement in the 20th century now creates vulnerability in the 21st: front-office work, once a gateway to middle-class employment, is now the occupation group most exposed to AI displacement.

4.3 The Entry Mechanism

Figure 2 visualizes the occupational distribution of each demographic group in AI exposure-complementarity space, where each bubble represents an occupation sized by its employment share within the group, and the black-outlined circle marks the group's employment-weighted average.

The contrast between the four panels captures the paper's central mechanism. Women entrants (top left) concentrate their employment mass in the HELC quadrant — high exposure, low complementarity — with the largest bubbles (shop assistants, clerks, cashiers) clustered

in the region where AI substitutes for human labor. Their weighted average sits squarely in the HELC zone. Men entrants (top right) distribute their mass more heavily toward the left side of the plot — the low-exposure region — reflecting concentration in manual and service occupations that AI affects only modestly. The weighted average for men falls near the LE–HELC boundary, substantially to the left of women's. Among incumbents, the pattern shifts: women incumbents (bottom left) show greater dispersion, with non-trivial employment mass in the HEHC quadrant — the professional and managerial occupations where AI augments productivity. Men incumbents (bottom right) remain concentrated in LE. The key visual takeaway is that gender, not age, determines vertical position (complementarity), while age determines horizontal spread:incumbents of both genders are more dispersed across the exposure axis than entrants, consistent with occupational resorting over the career.

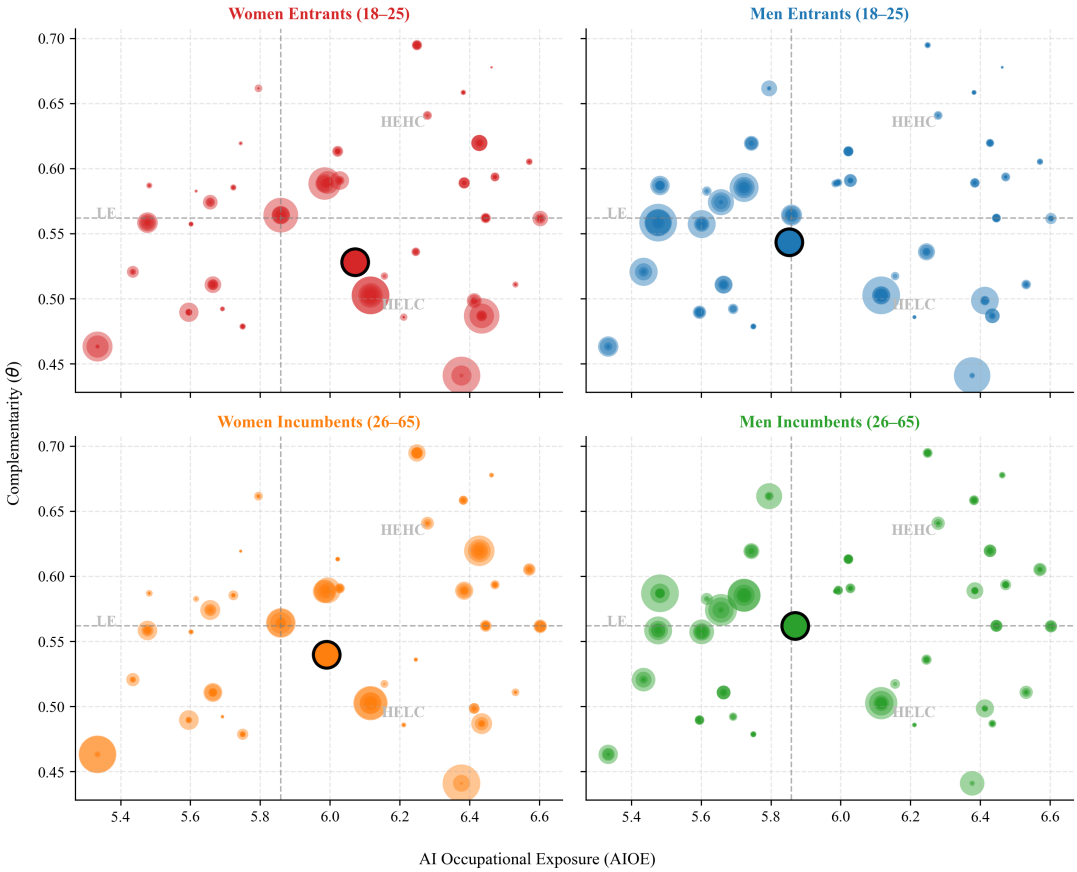


Figure 2: Occupational Distribution in AI Exposure-Complementarity Space by Demographic Group

Notes: Each bubble represents a 4-digit CBO occupation, sized by employment share within the group. Black-outlined circles denote employment-weighted group averages. Dashed lines indicate median splits defining the three AI risk categories: LE (low exposure), HEHC (high exposure, high complementarity), and HELC (high exposure, low complementarity). AIOE = AI Occupational Exposure Index (Felten et al., 2021); θ = complementarity score (Cazzaniga et al., 2024). PNADc 2023, all quarters pooled.

Since career stage appears to shape AI vulnerability — entrants are far more concentrated in

HELC than incumbents — a natural question is how the gender gap evolves over the lifecycle. Table 4 quantifies this using fine age bands.

Table 4: Gender Gap in HELC by Age

Age Group	Women (%)	Men (%)	Gap (p.p.)	W/M Ratio
18–20	57.7	28.6	29.2	2.02×
21–25	47.8	28.9	18.9	1.65×
26–30	40.5	25.4	15.1	1.59×
31–35	36.4	23.4	12.9	1.55×
36–40	31.9	21.3	10.5	1.50×
41–45	27.6	20.1	7.5	1.37×
46–50	24.7	19.3	5.3	1.28×
51–55	23.7	19.5	4.2	1.22×
56–60	23.5	20.3	3.2	1.16×
61–65	24.2	21.5	2.7	1.13×

Notes: HELC share calculated separately by gender and 5-year age band. Gap = Women − Men. Ratio = Women's HELC share divided by Men's HELC share.

Two patterns emerge from Table 4. First, the gender gap declines monotonically from 29.2 percentage points among the youngest workers (18–20) to 2.7 percentage points among the oldest (61–65). Women's HELC share falls steeply—from 57.7% to approximately 24%—while men's declines more gradually from 28.6% to approximately 21%. By ages 56–65, the gap has largely closed.

Second, the women-to-men ratio also declines monotonically—from 2.02× at ages 18–20 to 1.13× at ages 61–65. This indicates genuine convergence: not only does the absolute gap narrow, but relative exposure equalizes as well. The pattern is consistent with an entry mechanism—young women are channeled into HELC occupations at labor market entry and gradually transition out—rather than a lifecycle sorting story in which family responsibilities accumulate and push women into vulnerable positions.

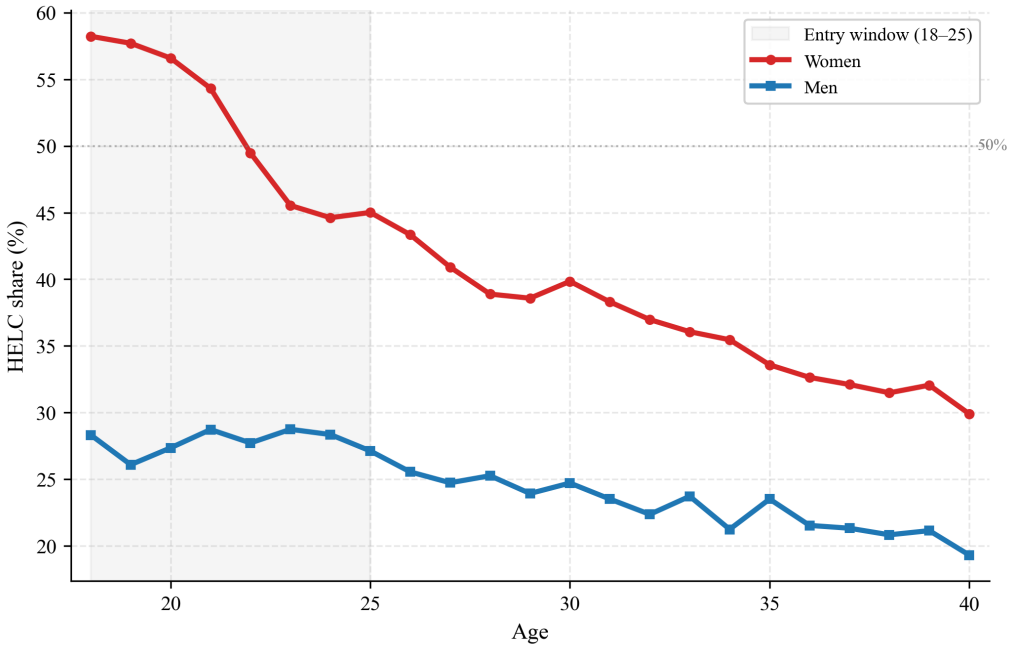


Figure 3: HELC Concentration by Age and Gender

Notes: Share of employed workers in High Exposure, Low Complementarity (HELC) occupations by single year of age and gender. PNADc 2023, all quarters pooled, weighted by V1028. Shaded area indicates the entrant cohort (18–25). The gap is widest at labor market entry and narrows with age as workers accumulate experience and reallocate across occupations.

Figure 3 traces the HELC share by age for women and men. The gap is widest at labor market entry (18–20) and narrows steadily with age, consistent with gradual occupational resorting as workers accumulate experience and credentials.

This convergence pattern, however, does not imply equal transition rates. As Section 4.5 documents, men exit HELC faster than women at every age. The convergence arises because women start from a much higher baseline: with 57.7% of young women in HELC versus 28.6% of young men, women have substantially more room to decline, and even a slower exit rate produces convergence in levels when the starting gap is large. The distinction between convergence in *levels* and divergence in *rates* is central to the theoretical model's calibration.

4.4 Heterogeneity by Education

Table 5 examines how the gender gap in HELC concentration varies by education level. The patterns are not monotonic and reveal an important interaction between gender and human capital.

Table 5: Gender Gap in HELC by Education Level

Education Level	Women (%)	Men (%)	Gap (p.p.)
No Schooling	19.3	9.8	+9.4
Incomplete Primary	13.8	9.5	+4.3
Complete Primary	19.7	12.8	+6.9
Incomplete Secondary	29.8	15.7	+14.1
Complete Secondary	41.3	24.4	**+16.9**
Incomplete Tertiary	48.7	43.5	+5.1
Complete Tertiary	32.5	36.6	−4.1

Notes: HELC share by education level and gender, pooling all ages.

Two patterns merit attention. First, the gender gap is largest among workers with complete secondary education (+16.9 p.p.), precisely the education level associated with pink-collar entry. Workers at this margin possess sufficient skills for office and sales work but typically do not proceed to professional occupations. Both HELC levels are elevated (41.3% for women, 24.4% for men), but women's disproportionate channeling into the pink-collar occupations documented in Table 3 generates a substantial gap.

Second, among workers with complete tertiary education, the gap reverses: college-educated women are 4.1 percentage points *less* likely than college-educated men to work in HELC occupations (32.5% vs. 36.6%). College education provides an alternative entry pathway through professional occupations that, while highly exposed to AI, exhibit high complementarity. The reversal does not imply protection, as HELC shares remain above 30% for both genders at the tertiary level, but reflects channeling into different types of high-exposure occupations.

These patterns illuminate the mechanism behind the pink-collar trap. Women with secondary education enter the labor market through sales and clerical positions that are uniformly classified as HELC—the same four occupations that account for 86% of the gender gap in AI vulnerability among young workers (Table 3). College education provides a partial escape route, redirecting women toward professional occupations with higher AI complementarity. The stability of this pattern [4] suggests a structural feature of Brazilian occupational segregation rather than a transient phenomenon.

This finding motivates the existence of an educational barrier in the model, as secondary education is sufficient for HELC but insufficient for HEHC jobs.

4.5 Occupational Transition Rates

The PNADc's rotating panel structure allows direct observation of occupational transitions, providing evidence on the dynamics underlying the cross-sectional patterns documented above. Table 6 presents quarterly transition rates by demographic group, based on 376,293 individual quarter-to-quarter transitions.

[4]The +16.9 percentage point gap at the secondary level is virtually identical to estimates from the 2010 Census.

Table 6: Quarterly Transition Rates: HELC Occupations

	Entrants (18–25)		Incumbents (26–65)	
	Women	Men	Women	Men
Panel A: Occupational Mobility				
HELC → HELC (stay)	89.2%	81.6%	85.1%	79.5%
HELC → HEHC (upward)	3.7%	4.3%	6.6%	8.3%
HELC → LE (lateral)	7.0%	14.0%	8.2%	12.2%
Total occupational exit	*10.8%*	*18.4%*	*14.9%*	*20.5%*
Panel B: Labor Force Exit				
HELC → Unemployment	4.8%	5.0%	2.1%	2.0%
HELC → Inactivity	**7.4%**	**5.3%**	**6.2%**	**2.8%**
Total labor force exit	*12.2%*	*10.3%*	*8.3%*	*4.8%*
Panel C: Inflow to HELC				
HEHC → HELC	13.1%	13.6%	8.6%	11.8%
LE → HELC	11.1%	7.3%	5.4%	4.2%
N (person-quarters in HELC)	10,184	8,001	40,193	36,942

Notes: Transition rates between consecutive quarters from PNADc 2023. Individual identifier constructed from household, person number, sex, and birth year. All rates weighted using survey calibration weights (V1028). Panels A–B: rates among workers initially in HELC. Panel C: rates among workers initially in HEHC or LE, respectively, who transition into HELC.

Three findings emerge from the transition analysis. First, young women have the lowest occupational exit rate from HELC: only 10.8% change occupation each quarter, compared to 18.4% for young men—a ratio of 1.71×. Among incumbents, the gap narrows but persists: men exit 1.38× faster than women (20.5% vs. 14.9%). Young women in pink-collar occupations are the most "stuck" of any demographic group.[5]

Second, Panel B reveals the key asymmetry that motivates the model's exit channel: unemployment is symmetric but inactivity is highly gendered. Young women and young men transition from HELC to unemployment at almost identical rates (4.8% vs. 5.0%), indicating that frictional displacement is gender-neutral. Inactivity transitions, however, diverge sharply: 7.4% of young women in HELC exit to inactivity each quarter, compared to 5.3% of young men. Among incumbents, the asymmetry is starker: women's inactivity rate is 2.21× that of men (6.2% vs. 2.8%). This pattern — symmetric unemployment, asymmetric inactivity — is consistent with gendered outside options driving voluntary exit rather than differential displacement, and corroborates the cross-country finding of Cazzaniga et al. (2025) that women in high-substitution occupations at the bottom of the wage distribution transition to inactiv-

[5]Among young women who do exit HELC, a larger share moves upward to HEHC (34% of exits) than among young men (23%), but since men exit at nearly twice the rate, absolute flows to HEHC are comparable. The salient difference is exit frequency, not exit quality.

ity more frequently. Section 4.6 investigates the household-level mechanisms underlying this asymmetry.

Third, Panel C shows that the labor market actively channels young women into HELC: 11.1% of young women in LE transition to HELC each quarter, compared to 7.3% for young men. This asymmetric inflow sustains the pink-collar concentration documented in Section 4.2 and implies that occupational reallocation away from HELC must work against a countercurrent of continued gender-differentiated entry.

These transition rates serve two roles in the theoretical framework. The occupational mobility rates (Panel A) calibrate the switching costs that govern how quickly workers can reallocate across sectors (Section 5.2). The inactivity transitions (Panel B) serve as an untargeted validation moment: the model's outside option structure, estimated independently from household characteristics in Section 4.6, should reproduce the gendered inactivity pattern without being calibrated to it. Section 5.5 confirms that it does.

4.6 What Drives Gendered Labor Force Exit?

The preceding analysis documents that transitions from HELC to unemployment are gender-neutral while transitions to inactivity are sharply gendered, ruling out differential displacement in favor of differential voluntary exit.

A natural question is whether this gender gap in exit behavior reflects intrinsic gender differences in labor market attachment, or whether it is explained by observable differences in household structure that raise women's reservation wages relative to men's. This distinction has direct implications for both modeling and policy: if household structure drives the exit gap, then the outside option can be parameterized from observable household characteristics and targeted by family policy; if a large residual gender effect remains, additional mechanisms must be invoked. To investigate this, I estimate a series of participation probits that progressively add household controls, exploiting individual-level variation in partner income and children within each demographic group. The central finding is that household structure explains more than half of the gender gap in labor force inactivity among entrants, with partner income and children operating as the primary channels.

To address how the household composition affects the labor force participation, I constructed three different household types: singles (S) who have no co-resident partner, Partnered without children (P_0), for those who live with a spouse or partner but have no household members under the age of 15, and Partnered with children (P_K), who lives with a partner and with at least one household member aged under 15. Partners are identified through the relationship-to-head (V2005): if the respondent is the head of the household, the partner is the co-resident spouse, and vice versa. For each partnered individual, I assign *partner income* as the partner's monthly labor income from the main job (VD4016), with partner income set to zero by construction when there is no partner (singles) or the partner is unemployed/inactive. Table 7 presents these distributions.

Table 7: Household Type Distribution by Demographic Group (%)

	Single (S)	Partnered, no kids (P0)	Partnered, w/ kids (PK)
WE: Women Entrants (18–25)	71.7	11.8	16.4
ME: Men Entrants (18–25)	84.0	7.6	8.4
WI: Women Incumbents (26–65)	40.3	30.5	29.2
MI: Men Incumbents (26–65)	35.6	31.5	32.9

Notes: PNADc 2023, all four quarters pooled, weighted by V1028. "Partnered" defined as co-resident spouse or partner identified through the household relationship variable (V2005). "Kids" defined as presence of household members aged under 15.

Most young entrants are single, particularly men (84%). Among the minority of young women who are partnered with children (16.4%), labor market behavior differs sharply, as the analysis below documents. Among incumbents, roughly 30% of women are partnered with children, creating a large population for whom household structure materially shapes labor force participation.

I estimate a weighted probit where the dependent variable is an indicator for being out of the labor force, on the full population of individuals aged 18–25 regardless of employment status (N = 210,531 person-quarters, weighted by V1028, standard errors clustered at the state level). I estimate three different specifications in sequence: the first includes only a Female indicator alongside age and state fixed effects; the second adds marital status, presence of children in the household, and log partner income; the third introduces the gender interactions to the preceding variables — this allows the household variables to operate differently for men and women. TThe average marginal effect of Female falls from 18.4 percentage points (spec 1) to 8.4 percentage points (spec 3): a reduction of 54.3%. This result reveals the importance of the household structure. Table 8 reports the full results.

Table 8: Determinants of Labor Force Inactivity: Entrants (18–25)

	(1) Baseline		(2) + Household		(3) Full	
	Coeff.	AME	Coeff.	AME	Coeff.	AME
Female	0.541***	18.4	0.517***	17.5	0.253***	8.4
	(0.040)		(0.038)		(0.023)	
Married			−0.254***	−8.6	−0.755***	−25.0
			(0.029)		(0.032)	
Has children (<15)			0.180***	6.1	−0.006	−0.2
			(0.016)		(0.014)	
Log partner income			0.026***	0.9	−0.048***	−1.6
			(0.008)		(0.009)	
Female × Married					1.390***	45.9
					(0.033)	
Female × Children					0.336***	11.1
					(0.031)	
Female × Partner inc.					−0.021***	−0.7
					(0.007)	
Years of schooling	−0.077***	−2.6	−0.075***	−2.5	−0.073***	−2.4
	(0.010)		(0.010)		(0.011)	
Age controls	Yes		Yes		Yes	
State FE	Yes		Yes		Yes	
N	210,531		210,531		210,531	
Pseudo R^2	0.069		0.072		0.092	
% Female explained	—		4.5%		**53.3%**	

Notes: Weighted probit, dependent variable = 1 if out of the labor force. Sample: all individuals aged 18–25 ($N = 210{,}531$). Standard errors clustered at UF level in parentheses. AME = average marginal effect in percentage points. The tertiary education variable is omitted for space (positive coefficient reflects enrollment; see footnote in text). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Three channels emerge: marriage (dominant), children (secondary), and partner income (small). Each of them is discussed next.

Marriage has opposite effects by gender. For young men, marriage reduces inactivity by 25.0 percentage points, consistent with the breadwinner role — marriage pushes men into the labor force. For young women, the net effect is reversed: marriage raises inactivity by 20.9 percentage points (−25.0 + 45.9). The Female × Married interaction of 45.9 percentage points is the single largest gendered predictor in the model. This pattern is consistent with household specialization: marriage creates a division of labor where women assume primary responsibility for household production, raising the reservation wage above the market wage. Critically, this effect operates through marital status itself rather than through partner income — conditional on being married, higher partner income slightly reduces inactivity for both genders.

Having children under 15 years-old raises young women's inactivity by 10.9 percentage points, but has no significant effect on young men. This asymmetry reflects the disproportionate allocation of caregiving responsibilities to women in the household. This implies that AI-

induced HELC wage declines will disproportionately affect mothers, for whom the opportunity cost of market work is highest.

When conditional on marital status, a higher partner's income slightly decreases inactivity for both genders (-1.6 p.p. for men, and -2.3 p.p. for women). The negative sign suggests that dual-earner couples are positively selected on labor market orientation. Education, by contrast, operates as consistent and gender-neutral buffer: each additional year of schooling reduces inactivity by 2.4 percentage points, confirming that human capital raises the market wage relative to the outside option. [6]

Table **??** reports the same analysis for incumbents (ages 26–65, $N = 1{,}010{,}834$). For this demographic group household structure explains 57.6% of the gender inactivity gap. The marriage differential narrows to 20.5 p.p. (vs 45.9 p.p. for entrants), but remains dominant. The children differential is closer for both groups (8.8 p. p. vs 11.1 p.p.). The qualitative finding — that marriage and children drive the gendered exit pattern through household specialization — is robust across career stages. [7].

Table 9 reports predicted inactivity rates from the full specification, evaluated at group-specific covariates means for each of the twelve group x household type cell.

[6] It is important to mention that the positive tertiary coefficient for entrants (+6.3 p.p.) reflects that many 18-25 years-old with university education are still enrolled and coded as inactive.

[7] As a supplementary check, I estimate the participation probit specifically for workers employed in HELC occupations using the PNADc quarterly panel (N=16,805 person-quarters). The coefficient on log wages is negative (0.366), implying that a 10% HELC wage decline raises quarterly exit probability by approximately 0.8 percentage points, but the limited sample yields imprecise estimates — no individual coefficient reaches conventional significance. Full results are reported in Appendix D

Table 9: Predicted Inactivity Rates by Demographic Group and Household Type (%)

	Single (S)	Partnered, no children (P0)	Partnered, with children (PK)
Entrants (18–25)			
Women (WE)	40.9	26.4	**59.6**
	[40.6, 41.3]	[25.6, 27.3]	[58.8, 60.3]
Men (ME)	28.1	4.2	5.8
	[27.8, 28.4]	[3.8, 4.7]	[5.3, 6.3]
Incumbents (26–65)			
Women (WI)	36.6	41.1	38.0
	[36.4, 36.8]	[40.8, 41.3]	[37.7, 38.3]
Men (MI)	20.0	13.1	**6.1**
	[19.8, 20.2]	[13.0, 13.3]	[6.0, 6.2]
Key contrasts:			
WE(PK) − ME(PK)			53.8 pp
WE(PK) − MI(PK)		53.4 pp	(ratio: 9.8×)
WE(PK) − MI(S)			39.6 pp

Notes: Predicted probabilities from the full-specification probit (column 3 of Tables 8–**??**) with group × household-type interactions, evaluated at cell-specific covariate means (age, education, state). 95% confidence intervals via the delta method in brackets. Bold cells mark the sharpest contrast: young partnered mothers (59.6%) versus incumbent partnered fathers (6.1%). $N = 1{,}221{,}365$.

For young men, partnership sharply reduces inactivity: predicted rates fall from 28.1% (single) to 4.2% (partnered, no children), and 5.8% (partnered with children), reinforcing the breadwinner role. For young women, partnership without children also reduces inactivity (from 40.9% to 26.4%), but children reverse this pattern: predicted inactivity for young partnered mothers reaches 59.6%. These findings suggest that it is not partnership per se, but motherhood within partnership that drives women's exit from the labor force. The most striking comparison arises between young partnered mothers and incumbent partnered fathers (59.6% vs 6.1%) — a nearly tenfold difference that reflects the full cumulative force of gender, age, and household specialization.

These findings inform the theoretical model framework in two ways: first, they identify the channels through which the outside option varies across groups: marital status and children account for more than half the gender exit differential, with marriage operating through household specialization rather than through partner income. The remaining 47% of the gap may reflect gender norms, preferences, or unobserved heterogeneity, which is absorbed by the idiosyncratic taste shock. Section 5 formalizes how these household-driven outside options interact with AI-induced sectoral wage changes to generate differential labor force exit across demographic groups.

5 Theoretical Framework

I develop a static general equilibrium model with heterogeneous workers, building on the framework of Ngai and Petrongolo (2017) for gendered sectoral sorting. Workers choose between three sectors and home production based on comparative advantage (Roy, 1951; Heckman and Sedlacek, 1985), with sector-specific work costs that depend on household structure following Goldin (2014). I extend this framework in three directions. First, I introduce AI-specific productivity shocks following Cazzaniga et al. (2024), which differentially affect sectors based on human-AI complementarity. Second, I add occupational switching costs calibrated from observed transition rates, following Traiberman (2019). Third, I incorporate educational barrier to high-complementarity occupations, generating a 'trap' mechanism for non-college workers.

I present the model in three stages: the environment (Section 5.1), the worker's optimization problem and equilibrium (Sections 5.2–5.3), and the estimation strategy that takes the model to data (Section 5.4)

5.1 Environment

5.1.1 Agents

The economy has four demographic groups $g \in \mathcal{G} = \{\text{WE}, \text{ME}, \text{WI}, \text{MI}\}$, where WE denotes women entrants (ages 18–25), ME men entrants, WI women incumbents (ages 26–65), and MI men incumbents.[8] Within each group there are three types of households $h \in \mathcal{H} = \{\text{S}, \text{P0}, \text{PK}$, where S denotes single, P0, partnered and with no children under the age of 15, and PK, partnered with children under the age of 15. The groups follow a distribution θ_{gh} observed directly in the data from the PNADc.

Each worker i in group g draws a productivity vector (a_{iH}, a_{iL}, a_{iE}) representing efficiency units in each sector, from a group-specific log-normal distribution: $\ln \mathbf{a}_i \sim \mathcal{N}(\boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g)$. The parameters μ_g and Σ_g are estimated from the selection-corrected wage distributions in Appendix A.

Each worker also has an outside option value v_i representing the utility flow from non-market activity:

$$v_i = \bar{v}_{gh} + \sigma_v \varepsilon_{i,v} \tag{1}$$

where $\bar{v}_{gh}$ is a group-by-household-type outside option value and $\varepsilon_{i,v}$ is an idiosyncratic taste shock. The twelve $\bar{v}_{gh}$ values are calibrated to match the predicted inactivity rates in Table 9, which range from 4.2% (partnered young men without children) to 59.6% (partnered young mothers). The participation analysis in Section 4.6 identified marriage and children as the dominant channels generating this variation; the model captures these effects through the household type structure without imposing that the coefficients are common across groups.

[8]The age×gender classification captures two empirically important sources of heterogeneity: gender differences in occupational sorting and life-cycle differences in labor market attachment. Section 3 documents both patterns in PNADc 2023 data.

5.1.2 Sectors and Technology

Three sectors $s \in \{\text{HEHC}, \text{HELC}, \text{LE}\}$ are classified by AI exposure and complementarity, where HEHC denotes High Exposure with High Complementarity, HELC is High Exposure with Low Complementarity, and LE is Low Exposure, following the measures developed by Felten et al. (2021) and Cazzaniga et al. (2024). Each sector produces a homogeneous final good using effective labor:

$$Y_s = A_s(\delta) \cdot L_s^{\beta} \tag{2}$$

where $A_s(\delta)$ is sector-specific total factor productivity depending on AI adoption intensity $\delta \in [0, 1]$, L_s is effective labor input aggregated in efficiency units, and $\beta \in (0, 1)$ captures decreasing returns, calibrated from observed sector employment shares and wage ratios (Section 5.4).

5.1.3 AI Productivity Shock

AI adoption affects each sector differentially through complementarity:

$$A_H(\delta) = A_H^0 \cdot (1 + \theta_H \kappa^+ \delta) \tag{3}$$

$$A_L(\delta) = A_L^0 \cdot (1 - (1 - \theta_L)\kappa^- \delta) \tag{4}$$

$$A_E(\delta) = A_E^0 \cdot (1 + \alpha_E \delta) \tag{5}$$

where θ_s is the complementarity score, $\kappa^+ = 0.30$ is augmentation intensity, $\kappa^- = 0.40$ is substitution intensity, and $\alpha_E = 0.02$ captures modest spillovers to low-exposure occupations.[9]

Under full adoption ($\delta = 1$), HEHC productivity rises by an average of 18.3%, HELC productivity falls by an average of 21.4%, and LE productivity rises by 2%, where the averages use $\bar{\theta}_s$, the employment-weighted mean complementarity within sector s. AI effects are sector-specific, not group-specific: all sector parameters are identical across demographic groups, and differential outcomes arise solely through occupational composition and exit behavior.

5.2 Worker's problem

Each worker i in group g with household type h chooses $d \in \{H, L, E, \text{Out}\}$ to maximize:

$$U_i(d) = \begin{cases} w_d \cdot a_{id} - c_{gd}(h) - \kappa_g \cdot \mathbf{1}[d \neq s_i^0] - B_H(e_i) \cdot \mathbf{1}[d = H] + \varepsilon_{id} & \text{if } d \in \{H, L, E\} \\ v_i + \varepsilon_{i,\text{Out}} & \text{if } d = \text{Out} \end{cases} \tag{6}$$

[9]The augmentation and substitution intensities ($\kappa^+ = 0.30$, $\kappa^- = 0.40$) are guided by the estimates in Eloundou et al. (2023) and the task-based framework of Acemoglu and Restrepo (2022), but are not directly estimated from Brazilian data. The asymmetry $\kappa^- > \kappa^+$ reflects the assumption that AI substitutes low-complementarity tasks more readily than it augments high-complementarity ones, consistent with the current generation of large language models performing routine cognitive tasks at near-human levels while augmenting complex judgment more modestly. The LE spillover $\alpha_E = 0.02$ is set conservatively. Appendix B reports sensitivity analysis varying κ^- from 0.20 to 0.60 and κ^+ from 0.15 to 0.45; the qualitative results—wage channel dominance and the WE–ME gap widening—are robust across the full range.

Where w_d is the endogenous wage per efficiency unit in sector d, determined in equilibrium. A worker with productivity a_{id} in sector d earns $w_d \cdot a_{id}$. $c_{gd}(h)$ is the cost of working in sector d for a worker of gender g and household type h. This captures non-pecuniary costs, such as children expenses, schedule rigidity, commuting, that vary by sector. Following Goldin (2014), occupations which are part of the HEHC sector are more inflexible (longer hours, travel requirements), imposing higher costs on workers with children. These costs are internally calibrated to match observed occupational distributions from section 5.4.3.

The switching costs κ_g applies when a worker moves from her baseline sectors s_i^0 — the one chosen before AI — to a different one. This captures the retraining costs, search frictions, and occupation-specific human capital loss. Following (Traiberman, 2019), switching costs are group-specific and calibrated from observed quarterly transition rates (table 6): groups with lower observed mobility face higher switching costs as young women exit HELC at 10.8% per quarter versus 18.4% for young men.

The educational barrier $B_H(e_i)$ restrict access to HEHC occupations. Workers without tertiary education ($e_i < \bar{e}_H = 15$ years of schooling) face a prohibitively high cost of entering professional and managerial occupations, reflecting credentialing requirements.[10] Following the logic of Hsieh et al. (2019), this barrier generates misallocation: workers who could be productive in HEHC are locked out by insufficient credentials. the barier is the mechanism that makes the "trap" binding — without it HELC workers could freely reallocate to HEHC when wages decline.

Lastly, $\varepsilon_{i,d}$ are i.i.d. Type 1 Extreme Value with scale σ_ε, yielding multinomial logit choice probabilities ((McFadden, 1974)). The probability that worker i chooses option d is:

$$P_i(d) = \frac{\exp\big(\bar{U}_i(d)/\sigma_\varepsilon\big)}{\sum_{d'} \exp\big(\bar{U}_i(d')/\sigma_\varepsilon\big)} \tag{7}$$

where $\bar{U}_i(d)$ is the deterministic component of utility — all terms in equation 6 except ε. The scale parameter σ_ε governs the elasticity of sector choice with respect to wage differentials and is calibrated to match the overall dispersion of occupational shares across groups.

5.3 Equilibrium

Aggregating individual choices over all four groups, household types, productivity draws, baseline sectors, and education levels yields effective labor supply in each sector:

$$L_s(\mathbf{w}) = \sum_{g \in \mathcal{G}} \eta_g \sum_{h \in \mathcal{H}} \phi_{gh} \sum_{s^0} \pi^0_{gs^0} \int \int a_{is} \cdot P_i(s \mid \mathbf{a}, e_i, h, s^0, g; \mathbf{w})\, dF_g(\mathbf{a})\, dG_g(e) \tag{8}$$

where η_g is the population share of group g, ϕ_{gh} is the share of household type h within group g, $\pi^0_{gs^0}$ is the baseline (pre-AI) share of group g in sector s^0, $F_g(\mathbf{a})$ is the group-specific productivity distribution, $G_g(e)$ is the group-specific education distribution, and $P_i(s \mid \cdot)$ is the

[10] The entrant group (18–25) includes workers who have not yet completed tertiary education but may do so in the future. In the static model, the barrier captures *current* credential constraints: an 18-year-old without a degree is locked out of HEHC regardless of whether the cause is age or educational choice. Appendix C reports results restricting entrants to ages 22–26, for whom the absence of tertiary credentials more plausibly reflects access rather than timing. The qualitative findings are robust, with attenuated magnitudes reflecting lower HELC concentration in the older cohort.

multinomial logit choice probability from equation (6). The choice probability depends on the worker's baseline sector s^0 through the switching cost $\kappa_g \cdot \mathbf{1}[s \neq s^0]$ and on education e_i through the barrier $B_H(e_i) \cdot \mathbf{1}[s = H]$. In the frictionless specification (v1), the dependence on s^0 and e_i vanishes and the expression simplifies to a standard Roy-model aggregation.

Wages are determined by the marginal product of effective labor:

$$w_s = \beta \, A_s(\delta) \, L_s^{\beta-1} \tag{9}$$

With $\beta < 1$, the wage is decreasing in sector employment: sectors that attract more workers pay lower wages per efficiency unit, reflecting diminishing returns.

A competitive equilibrium is a wage vector $\mathbf{w}^* = (w_H^*, w_L^*, w_E^*)$ such that equations (8) and (9) hold simultaneously in all three sectors. This is a system of three equations in three unknowns, solved numerically by dampened fixed-point iteration.[11]

A key property of the general equilibrium is the *wage buffer*: when AI reduces HELC productivity and workers exit, the decline in L_L partially restores w_L through decreasing returns—the remaining HELC workers become scarcer and therefore better paid. Symmetrically, as displaced workers flow into HEHC and LE, the increase in L_H and L_E dampens wage gains in those sectors. General equilibrium thus moderates the AI shock in both directions relative to a partial equilibrium analysis that holds wages fixed at $w_s^{\text{PE}} = w_s^0 \cdot A_s(\delta)/A_s^0$.

The model endogenously determines five sets of outcomes: equilibrium wages $\mathbf{w}^*$, occupational distributions π_{gs} (the share of group g in sector s), labor force participation rates ℓ_g, average earnings by group, and the gender exit ratio $\rho = \xi_W/\xi_M$—the relative responsiveness of female and male participation to HELC wage changes.

5.4 Estimation and Calibration

5.4.1 Comparative Advantage Distributions

Observed wages are selected, as we observe each worker only in the sector she chose. Thus, estimating productivity distribution from the observed wages without correcting for selection gives biased means. I estimate a selection-corrected wage distributions using a two-step procedure following Lee (1983) and Dahl (2002). The first step estimates a multinomial logit for sector choice on observables (education, age, region, household type) separately for each group, yielding predicted probabilities. The second step uses these probabilities to construct a control function that corrects for selection in sector-specific wage regressions. The corrected productivity distribution reveal that women's relative productivity in HELC exceeds men's, what suggests

[11] The algorithm starts from observed wages, computes group-level choice probabilities using a Berry (1994) inversion of the multinomial logit, derives implied wages from equation (9), and updates using $\mathbf{w}^{(k+1)} = 0.3\,\mathbf{w}^{\text{implied}} + 0.7\,\mathbf{w}^{(k)}$. Convergence is achieved when the maximum relative wage change falls below 10^{-4}, typically within 50–200 iterations. The computational implementation aggregates group-level employment shares rather than evaluating the efficiency-unit integral in equation (8) directly: mean productivity is absorbed into the calibrated A_s^0, and 400,000 simulated workers are used to compute individual-level choice probabilities which are then aggregated to sectoral shares. The implementation uses a log-linearized Berry-style formulation $\omega \ln w_s$ at the group level, a standard simplification in quantitative GE models (see e.g., Traiberman, 2019). Since all parameters are jointly calibrated to match observed wages and employment shares, the baseline equilibrium is exact regardless of functional form. Individual-level earnings are computed using the full productivity distributions from Block 1. See Appendix F for details.

that women are in HELC not because they lack skills for other sectors but because comparative advantage and structural constraints channel them there. The full estimation details, first-stage diagnostics, and the corrected wage distributions for all groups are reported in Appendix A.

5.4.2 Outside Option Parameters

The outside option parameters are taken directly from the participation analysis in section 6.1. Marital status and children were identified as the dominant channels, with marriage generating a 45.9 percentage points gender reversal and children adding extra 11.1 p.p. differential. The predicted inactivity rates from Table 9 — which range from 4.2% to 59.6% (from partnered young men without children vs partnered young mothers) – serve as the calibration targets for the outside option values $\bar{v}_{g,h}$ in the model.

5.4.3 Internal Calibration

The sector-specific work costs $c_{gd}(h)$ are calibrated internally to reproduce the observed occupation distributions from Table 1: 12 sector x group shares plus 4 LFP rates provide 16 moments for calibration. The switching costs κ_g are calibrated from the observed quarterly transition rates in Table 6 so that higher observed stickiness (which means lower exit rates from a sector) implies higher switching costs. The logit scale parameter σ_ε is set to match the overall dispersion of occupational shares across groups. Lastly, the baseline TFP levels A_s^0 and the decreasing returns parameter β are calibrated to to match observed sector employment shares and wage ratios from PNADc 2023.

5.5 Baseline validation

The estimated model should match targeted moments by construction (those being occupational shares, LFP, and wages), however the key test relies on whether it also matches untargeted moments. The model generates a gender exit ratio $\rho = 2.11$, closely matching the 2.08 observed PNADc transition data. Exit rates by household type match the empirical 12-cell table: the model predicts 44.4% inactivity for partnered mothers versus the observed 43.3%. Sector shared match within 0.3 percentage points for all group x sector cells.

The close fit on untargeted moments confirm that the structural model estimated from wage distributions and participation probits correctly captures the behavioral mechanism generating occupational sorting and differential exit. Table 10 shows all model vs data comparisons.Figure 4 plots model-predicted against observed moments. Targeted moments (occupational shares, LFP rates) cluster tightly along the 45-degree line by construction. The untargeted moments — the gender exit ratio and household-type-specific inactivity rates — also fall close to the line, confirming that the structural model captures the behavioral mechanisms generating differential exit without being calibrated to these moments directly.

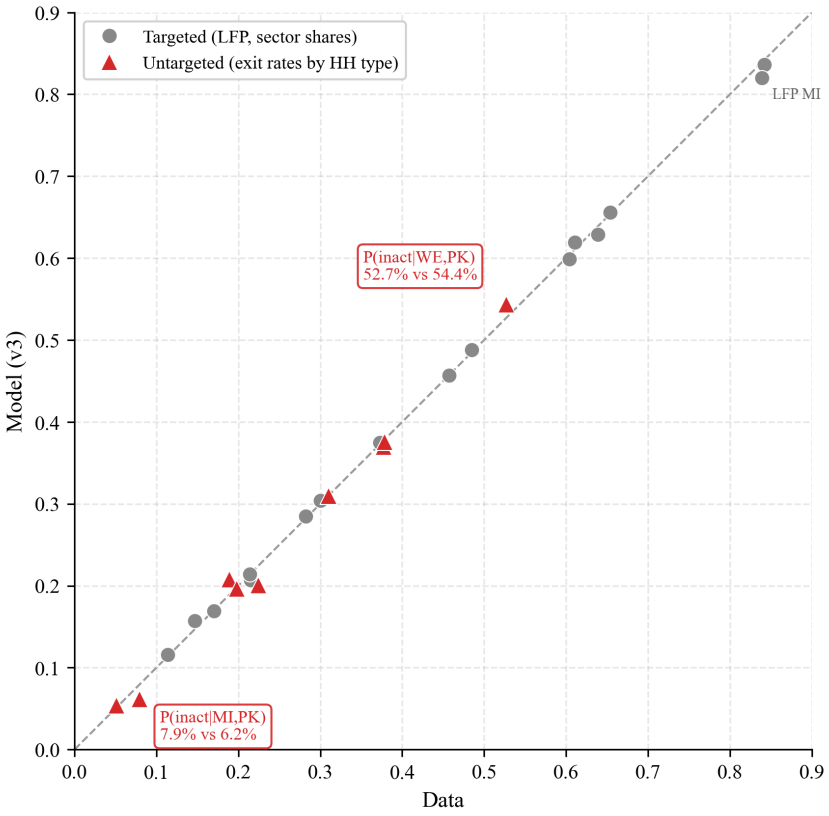


Figure 4: Baseline Validation: Model vs. Data Moments

Notes: Each point represents a moment from the model (vertical axis) plotted against its empirical counterpart (horizontal axis). Circles denote targeted moments (occupational shares, labor force participation rates); triangles denote untargeted moments (gender exit ratio ρ, inactivity rates by household type). The dashed line indicates perfect fit. Targeted moments match by construction; the close fit of untargeted moments provides an overidentification test of the structural model.

Table 10: Baseline Validation: Model vs. Data

Moment	Data	Model	Targeted?
Occupational Shares (%)			
WE in HELC	50.5	50.2	Yes
ME in HELC	28.8	28.9	Yes
WI in HEHC	21.5	21.3	Yes
MI in LE	63.7	63.5	Yes
Labor Force Participation (%)			
WE	45.7	45.9	Yes
ME	61.5	61.3	Yes
WI	53.5	55.1	Yes
MI	75.5	75.3	Yes
Untargeted Moments			
Gender exit ratio ρ	2.08	2.11	No
P(inact | WE, PK)	43.3%	44.4%	No
P(inact | MI, PK)	6.5%	6.1%	No

Notes: Data moments from PNADc 2023. Model moments from the baseline equilibrium ($\delta = 0$). The gender exit ratio $\rho = \xi_W/\xi_M$ and household-type-specific inactivity rates were not used in calibration.

Crucially, ρ was not targeted in the calibration — it emerges from the estimated productivity distributions and outside options, providing an overidentification test of the structural model.

5.6 The Three Versions

I solve the model under three friction specifications to trace how labor market rigidities amplify AI's gender impact. Table 11 summarizes the three versions.

Table 11: Model Versions

Version	κ_g	$B_H(e_i)$	Choice set	Interpretation
V1: Frictionless	0	0	$\{H, L, E, \text{Out}\}$	Long-run full adjustment
V2: Stickiness	> 0	0	$\{H, L, E, \text{Out}\}$	Occupational frictions
V3: Sticky + barrier	> 0	> 0	$\{L, E, \text{Out}\}^*$	Frictions + credential gate

Notes: *For workers without tertiary education ($e_i < 15$). Workers with tertiary education retain access to HEHC. Version 3 is the preferred specification.

The progression from V1 to V3 traces how frictions transform the AI shock's gender impact. In V1, workers freely reallocate and the gap increase is modest. Adding switching costs (V2) slows adjustment and traps workers in declining sectors. Adding the educational barrier (V3)

blocks the upward escape route to HEHC for non-tertiary workers, making the trap binding. Section 6 presents results for all three versions.

5.7 Decomposition

I decompose the change in the earnings gap into three channels:

$$\Delta\text{EarnGap} = \underbrace{\Delta^{\text{Wage}}}_{\text{direct shock}} + \underbrace{\Delta^{\text{Realloc}}}_{\text{reallocation}} + \underbrace{\Delta^{\text{Exit}}}_{\text{labor force exit}} \tag{10}$$

defined as:

$$\Delta^{\text{Wage}} = \text{Gap}(\boldsymbol{\pi}^0, \mathbf{w}^*, \boldsymbol{\ell}^0) - \text{Gap}(\boldsymbol{\pi}^0, \mathbf{w}^0, \boldsymbol{\ell}^0) \tag{11}$$

$$\Delta^{\text{Realloc}} = \text{Gap}(\boldsymbol{\pi}^*, \mathbf{w}^*, \boldsymbol{\ell}^0) - \text{Gap}(\boldsymbol{\pi}^0, \mathbf{w}^*, \boldsymbol{\ell}^0) \tag{12}$$

$$\Delta^{\text{Exit}} = \text{Gap}(\boldsymbol{\pi}^*, \mathbf{w}^*, \boldsymbol{\ell}^*) - \text{Gap}(\boldsymbol{\pi}^*, \mathbf{w}^*, \boldsymbol{\ell}^0) \tag{13}$$

where superscript 0 denotes baseline (pre-AI) values and $*$ denotes post-AI equilibrium values. The wage effect captures the immediate impact of differential AI productivity shocks, holding occupational composition and participation fixed. The reallocation effect captures compositional upgrading as workers move across sectors, holding participation fixed. The exit effect captures differential labor force withdrawal—the residual after accounting for wages and reallocation.

6 Quantitative Analysis

This section presents the quantitative results from the calibrated model. I organize the analysis in four parts: the main impact of AI on the gender earnings gap across demographic comparisons (Section 6.1), the three-way decomposition into wage, reallocation, and exit channels (Section 6.2), the role of labor market frictions in amplifying the effect (Section 6.2.1), and policy counterfactuals (Section 6.3).

6.1 Main results

I compute four earnings gap measures to capture the heterogeneity of AI's gender impact. The earnings gap between groups g and g' is defined as:

$$\text{Gap}_{g,g'} = \frac{\text{Earn}_{g'} - \text{Earn}_g}{\text{Earn}_{g'}} \times 100 \tag{14}$$

where $\text{Earn}_g = \bar{w}_g \times \ell_g$ is group g's average wage multiplied by its labor force participation rate. Table 12 reports the change in this gap under full AI adoption for four comparisons.

Table 12: Effect of AI on Gender Earnings Gap: Preferred Specification (v3)

Comparison	What it captures	Δ gap (pp)	Interpretation
WE vs ME	Pure gender (same cohort)	$+2.43$	Occupational segregation at entry
WE vs MI	Gender + career stage	$+3.44$	Maximum exposure differential
WI vs MI	Gender among incumbents	-0.69	HEHC women benefit from AI
All W vs M	Aggregate	-0.64	Entry and incumbent effects cancel

Notes: Change in earnings gap under full AI adoption ($\delta = 1$) relative to baseline ($\delta = 0$). Positive values indicate gap widening (women worse off relative to men). Earnings defined as mean wage $\times$ labor force participation rate. Version 3: switching costs + educational barrier.

AI widens the gender earnings gap among labor market entrants by 2.43 percentage points. This is the pure gender effect within the same age cohort — comparable career stage, similar education levels — where the only systematic difference is occupational sorting. The gap between entrant women and incumbent men widens further, by 3.44 percentage points, reflecting both occupational segregation and the career-stage advantage of incumbent men, who benefit from higher HEHC representation. Among incumbents, the gap narrows by 0.69 percentage points: women incumbents hold 21.5% of HEHC jobs, enough to benefit from AI-driven productivity gains in professional occupations. The aggregate all-women-vs-all-men gap shows a similar magnitude (-0.64 pp) but for a different reason: the losses of young entrants and the gains of incumbent women nearly cancel. Any analysis that pools age groups will miss the entrant-specific vulnerability entirely — this aggregation artifact is a methodological warning for the growing literature on AI and gender inequality.

6.2 Three-way decomposition

Table 13 decomposes the gap change into three channels defined in section Section 5.7.

Table 13: Three-Way Decomposition of Earnings Gap Change (v3)

Channel	WE vs ME		WE vs MI		WI vs MI		All W vs M	
	Δ	Share	Δ	Share	Δ	Share	Δ	Share
Wage	$+2.03$	83%	$+3.10$	90%	-0.60	87%	-0.56	88%
Reallocation	-0.49	-20%	-0.23	-7%	-0.44	64%	-0.44	69%
Exit	$+0.90$	37%	$+0.57$	16%	$+0.35$	-51%	$+0.36$	-57%
Total	$+2.43$	100%	$+3.44$	100%	-0.69	100%	-0.64	100%

Notes: Decomposition follows equations (11)–(13). Δ in percentage points. Shares may not sum to 100% due to rounding. Positive values indicate gap widening. For WI vs MI and All W vs M, the total change is negative (gap narrows), so the exit channel's positive sign indicates it partially offsets the narrowing.

The wage channel dominates across all comparisons, accounting for 83–90% of the gap widening for entrant women. HELC wages fall by approximately 20% while HEHC wages rise

by 16%. Young women, with 50.5% HELC concentration, absorb the decline while young men, more evenly distributed, benefit from HEHC gains. The exit channel contributes +0.57 p.p. for WE vs MI (which represents 16% of the widening), and 0.90 p.p. for WE vs ME (which accounts for 37% of the widening). The larger exit share in the within-cohort comparison reflects that young women's higher outside options which are driven by marriage and children (as suggested in Section 4.6)). These generate a sharper exit differential relative to young men of the same age, who have low baseline inactivity rates regardless of household structure.

The reallocation channel provides a modest buffer, reducing the gap widening by 0.49 p.p. in the same cohort, and 0.23 p.p. for entrants vs incumbents. some women switch from HELC to HEHC and LE, partially offsetting the wage effect, but the existence of the educational barrier limits this channel to the minority of HELC workers with the tertiary education credentials. For incumbents and in the aggregate, the same channels operate in a different balance. The wage effect narrows the gap (incumbent women benefit from HEHC), reallocation narrows it further, but the exit channel partially offsets the narrowing — even among incumbents, women exit at higher rates than men when sectoral wages shift.

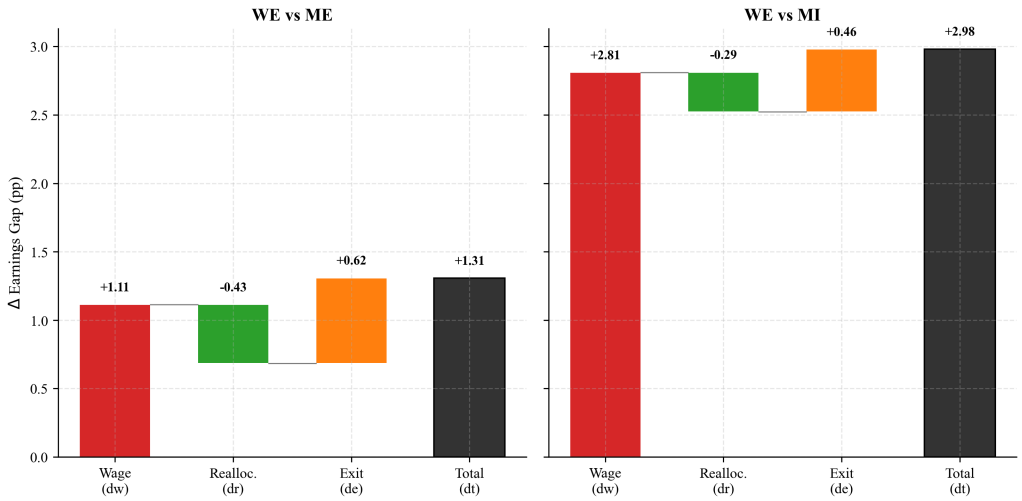


Figure 5: Three-Way Decomposition of AI-Induced Earnings Gap Change

Notes: Waterfall decomposition of the change in the earnings gap under full AI adoption ($\delta = 1$), preferred specification (v3). Left panel: WE vs ME (pure gender, same cohort). Right panel: WE vs MI (gender + career stage). Each bar shows the incremental contribution of one channel. The wage channel (red) dominates in both comparisons. The reallocation channel (green) provides a modest buffer. The exit channel (orange) adds a secondary contribution, larger for the within-cohort comparison.

Figure 5 presents the decomposition visually. Two patterns stand out. First, the wage channel accounts for the bulk of the gap widening in both comparisons, confirming that the primary harm operates through depressed HELC earnings rather than labor force withdrawal. Second, the exit channel's contribution is larger for WE vs ME (37%) than for WE vs MI (16%), reflecting that the household specialization mechanism documented in Section 4.6 generates sharper exit differentials within the entry cohort than across cohorts. This distinction has direct implications for the policy counterfactuals in Section 6.3: interventions that address wages (retraining,

occupational reallocation) target the dominant channel, while childcare subsidies target the secondary but non-negligible exit channel.

6.2.1 The role of frictions

Table 14 traces how the gap widening changes as labor market frictions are progressively introduced to the model.

Table 14: Earnings Gap Change by Model Version (pp)

Version	WE vs MI	WE vs ME	WI vs MI	All W vs M
v1: Frictionless	$+2.96$	$+1.17$	-1.07	-0.87
v2: Stickiness	$+3.32$	$+2.04$	-0.81	-0.56
v3: Sticky + barrier	$+3.44$	$+2.43$	-0.69	-0.40

Notes: Positive values indicate gap widening. v1: $\kappa_g = 0$, $B_H = 0$. v2: $\kappa_g > 0$, $B_H = 0$. v3: $\kappa_g > 0$, $B_H > 0$ (preferred). All under full AI adoption ($\delta = 1$).

For the within-cohort comparison, the gap widening more than doubles from +1.17 p.p. (frictionless) to +2.43 p.p. (full specification). Each friction amplifies the effect by reducing young women's ability to escape declining HELC occupations. Adding switching costS raises the WE vs ME gap widening by 0.87 p.p.. Young women's higher occupational stickiness [12] slows their reallocation toward rising-wage sectors, trapping them longer in the HELC as wages decline. Adding the educational barrier, contributes to a further 0.39 p.p. increase, by blocking upward escape route to HEHC for workers without tertiary credentials (about 80% of young women in HELC) however, frictions amplify the gap widening rather than creating it: even in frictionless model, the WE vd ME gap widens by 1.17 p.p., as the fundamental mechanism is occupational segregation interacting withe differential AI productivity shocks. Frictions determine how much of the vulnerability translates into persistent disadvantage.

6.3 Counterfactuals

The model's structure allows counterfactual experiments evaluating how specific interventuons would alter AI's gender impact. I consider two classes: education policies that expand access to HEHC, and distribution policies that reduce occupational segregation at entry. Table 15 reports the pre-AI baseline gap an the AI-induced change under each scenario. Each counterfactual affects the earning gap through two margins; the baseline (pre-AI) gap, which changes because the policy alters occupational sorting itself, and the AI-induced change, which captures how the policy modifies vulnerability to the technology shock.

[12] As shown in 4.5 the quarterly exit rate from HELC women is 10.8% versus 18.4% for young men, both stable in all the evaluated periods.

Table 15: Policy Counterfactuals: Earnings Gap Levels and AI-Induced Changes

	WE vs MI		WE vs ME		WI vs MI		All W vs M	
Policy	Base	Δ	Base	Δ	Base	Δ	Base	Δ
Benchmark								
V3: Sticky + barrier	59.6	+3.44	23.0	+2.43	32.9	−0.69	32.5	−0.40
Education counterfactuals								
1A: Remove barrier	61.7	+2.57	18.5	+3.10	33.4	+0.70	31.6	+0.92
1B: Retrain 35% WE	50.5	+0.26	6.0	−4.95	33.0	−0.68	29.9	−0.98
1C: Preventive education	56.1	+2.24	16.4	−0.37	33.0	−0.69	30.9	−0.62
1D: Lower barrier (12 yrs)	57.5	+1.56	18.6	+1.75	33.1	−0.09	31.2	+0.09
Distribution counterfactuals								
2A: Desegregate (WE→ME)	50.5	+4.18	5.8	+2.93	32.9	−0.69	29.9	−0.27
2B: Reverse (ME→WE)	59.6	+3.44	6.0	+2.78	32.9	−0.69	30.2	−0.35
2C: Top-occ equalize	51.3	+4.36	7.2	+3.37	32.9	−0.69	30.0	−0.24
Combined								
3C: Retrain + Desegregate	41.2	+0.71	−11.8	−5.09	33.0	−0.68	28.2	−0.91

Notes: "Base" is the pre-AI earnings gap (%) under each policy scenario; "Δ" is the AI-induced change (pp, $\delta = 1$). Positive Δ indicates gap widening. Baseline gaps differ across policies because counterfactuals alter the pre-AI sectoral allocation. See text for policy descriptions.

Targeted retraining is the most effective single intervention. Providing tertiary-equivalent credentials to 35% of non-degree young women in HELC reduces the baseline WE vs ME gap from 23.0% to 6.0% and reverses the AI-induced change from +2.43% to -4.95%. Retrained women do not merely survive AI shock, they benefit from it, as HEHC access exposes them to AI-augmented productivity gains. [13] Less intensive education interventions also improve outcomes but more modestly. Preventive education (giving young women the education distribution of incumbent men – 1C) reduces the baseline gap to 16.4% and neraly eliminates the AI-inuced widening (-0.37p.p.). Lowering the barrier threshold to 12 years of schooling (1D), equivalent to techinical certification, brings the baseline gap to 18.6% and reduces AI-induced widening to +1.75p.p.. Simply removing the barrier without providing credentials (1A) has smaller effects and paradoxically increases AI-induced widening to +3.10p.p., as newly mobile workers compete for HEHC positions, compressing wages through the GE channel.

Distribution counterfactuals operate through a different margin. Full desegregation — giving young women the occupational distribution of young men (2A) — dramatically reduces the baseline gap from 23.0% to 5.8%, nearly eliminating pre-existing inequality. However the AI-induced widening slightly increases to +2.93p.p., because men's occupational distribution

[13] Counterfactual 1B asks what would happen if 35% of non-degree young women currently in HELC held tertiary-equivalent credentials. In practice, achieving this would require a combination of instruments: retraining programs for workers already in the labor market (ages 22 and above), and expanded access to higher education for those still in the educational pipeline (ages 18–21). The model does not distinguish between these delivery mechanisms —it evaluates the equilibrium effect of the credential itself.

concentrates workers in LE (62.6%), where AI productivity gains are modest. Desegregation addresses the starting point: where women work before AI, but does not change how they respond to the shock.

Targeting only four occupations (2C): shop assistants, clerks, cashiers, and receptionists (86% of the HELC concentration gap) , captures most of the desegregation effect on the baseline gap (7.2% vs 5.8%). Focused interventions on small number of specific occupations could substantially reduce gender inequality without requiring comprehensive labor market reform. These results reveal a fundamental distinction. Educational policies change the response to AI by unlocking the access to HEHC sector and turning the shock less harmfull and more beneficial for retrained workers. Distribution policies change the starting point by reducing occupational segregation, shrinking the pre-existing gap that AI amplifies. The most effective strategy combines both margins.

Combining retraining with desegregation (3C) produces the largest effect: the baseline WE vs ME gap reverses to -11.8% (young women earn more than young men before AI), and AI reinforces this advantage ($\Delta = -5.09\,\text{pp}$), as retrained women in desegregated occupations benefit from AI-augmented productivity in HEHC occupations. This represents an upper bound of what comprehensive labor market reform could achieve.

All counterfactuals leave incumbent gap essentially unchanged ($\tilde{3}3\%$), confirming that interventions targeting entrant women's occupational sorting do not affect incumbent women who have already established their sectoral position — different cohorts require different policy approaches.

6.3.1 Decomposition Results

6.4 Sensitivity Analysis

I examine the robustness of the main findings to four key parameters: the decreasing returns parameter β, the logit scale σ_ε, the HELC substitution intensity κ^-, and the children effect on outside options α_2. The qualitative conclusions — wage channel dominance, the WE–ME gap widening, and the aggregation artifact — are robust across the full parameter range (Appendix B). The magnitude of the gap widening is most sensitive to κ^-, which directly governs the HELC productivity decline: varying κ^- from 0.20 to 0.60 changes the WE vs ME gap widening from approximately +1.0 to +4.0 pp, with the wage channel share remaining above 80% throughout. Results are also robust to redefining the entry cohort as ages 22–26 (Appendix C): the effects attenuate, as expected given lower HELC concentration in the older cohort, but all qualitative findings are preserved.

7 Conclusion

This paper examines how AI adoption affects gender earnings inequality through pre-existing occupational segregation. Combining descriptive evidence and participation analysis from PNADc microdata with a general equilibrium model featuring comparative advantage sorting, occupational switching costs, and educational barriers, I quantify the chanlles through which AI differentially affects men and women across different age life cycle moments. The central finding

is that AI widens the gender earning gap among labor market entrants by 2.4 percentage points, with the wage channel accounting for 83% of the effect. Young women, half of whom work in AI-vulnerable clerical and sales occupations, absorb declining HELC wages while young men, distributed across less-exposed sectors, are largely isolated. Aggregate gender comparisons fail to detect this effect because incumbent women in professional occupations benefit from AI complementarity, offsetting the losses of young entrants.

Two empirical findings underpin the model's quantitative results. First, household structure — marriage and children — explain more than two thirds of the gender gap in labor force exit, with marriage generating reversal of 36 percentage points between men an women. The mechanism is household specialization upon marriage, not spousal insurance through partner income. Second, labor market frictions amplify AI's gender impact: switching costs and the educational barrier together double within-cohort gap, widening from 1.2 to 2.4 percentage points. The educational barrier is what makes the pink-collar trap binding, as it blocks the upward escape route to AI-augmented professional occupations for the approximately 85% of young women in HELC who lack tertiary education. Witout it, workers could freely reallocate to high-complementarity sectors, and the gap widening would be substantially smaller.

The counterfactual experiments points to education as the critical policy lever. Retraining 35% of non-degree women in vulnerable occupations eliminates the AI-induced gap widening entirely, and retrained women do not merely survive the shock, but actually benefit from it by accessing HEHC AI-augmented productivity occupations. Occupational desegregation targeting four specific occupations (shop assistants, clerks, cashiers, and receptionists) captures most of this effect without requiring comprehensive labor market reform. More broadly, the reults reveal a fundamental distinction between two classes of intervention: education policies change how women respond to AI by unlocking access to complementary occupations, while distribution policies change were women start by reducing the pre-existing segregation. The most effective policy combines both margins.

Several limitations deserves to be mentioned. The model is static and cannot capture the dynamic adjustment path, anticipatory educational choices, or scarring effects from early-career displacement. The AI shock magnitudes are drawn from international estimates rather than estimated from Brazilian data, and the degree of occupational segregation, the educational barrier, and the household specialization patterns documented here may differ in other contexts. Future work could extend the analysis in three directions: incorporating dynamics with adjustment costs to trace the transition path, estimating AI productivity effects directly from emerging Brazilian data as AI adoption progresses (using RAIS for instance), and conducting a cross-country comparisons to assess whether the aggregated gap lack of effect and the cohort stronger effects is a feature of brazil's labor market or a more general phenomenon.

References

Acemoglu, D. and Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542.

Acemoglu, D. and Restrepo, P. (2022). Tasks, automation, and the rise in U.S. wage inequality. *Econometrica*, 90(5):1973–2016.

Aguiar, M., Bils, M., Charles, K.K., and Hurst, E. (2017). Leisure luxuries and the labor supply of young men. *NBER Working Paper* 23552.

Autor, D.H. (2019). Work of the past, work of the future. *AEA Papers and Proceedings*, 109:1–32.

Autor, D.H., Levy, F., and Murnane, R.J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

Autor, D.H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3):3–30.

Autor, D.H. (2022). The labor market impacts of technological change: From unbridled enthusiasm to qualified optimism to vast uncertainty. *NBER Working Paper* 30074.

Bargain, O., Orsini, K., and Peichl, A. (2014). Comparing labor supply elasticities in Europe and the United States: New results. *Journal of Human Resources*, 49(3):723–838.

Beaudry, P., Green, D.A., and Sand, B.M. (2016). The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics*, 34(S1):S199–S247.

Black, S.E. and Spitz-Oener, A. (2008). Explaining women's success: Technological change and the skill content of women's work. *Review of Economics and Statistics*, 92(1):187–194.

Blau, F.D. and Kahn, L.M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In Ashenfelter, O. and Card, D. (eds.), *Handbook of Labor Economics*, Vol. 3A, pp. 1559–1695. Elsevier.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Brussevich, M., Dabla-Norris, E., Kamber, G., and Kochhar, K. (2018). Gender, technology, and the future of work. *IMF Staff Discussion Note* 2018/07.

Brynjolfsson, E. and McAfee, A. (2014). *The Second Machine Age*. W.W. Norton & Company.

Budig, M.J. and England, P. (2001). The wage penalty for motherhood. *American Sociological Review*, 66(2):204–225.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and labor market dynamics: General equilibrium analysis of the China trade shock. *Econometrica*, 87(3):741–835.

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A.J., Pizzinelli, C., Rockall, E., and Tavares, M.M. (2024). Gen-AI: Artificial intelligence and the future of work. *IMF Staff Discussion Note* 2024/001.

Cazzaniga, M., Pizzinelli, C., Rockall, E., and Tavares, M. M. (2024). Exposure to artificial intelligence and occupational mobility: A cross-country analysis. IMF Working Paper 2024/116.

Cazzaniga, M., Panton, A., Li, L., Pizzinelli, C., and Tavares, M. M. (2025). A gender lens on labor market exposure to AI. *AEA Papers and Proceedings*, 115:56–61.

Cerina, F., Moro, A., and Rendall, M. (2021). World trade and the gender wage gap. *Journal of International Economics*, 131:103464.

Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records. *The Quarterly Journal of Economics*, 126(2):749–804.

Cortés, P. and Pan, J. (2018). Occupation and gender. In *The Oxford Handbook of Women and the Economy*, 425–452.

Cortés, P. and Pan, J. (2020). Children and the remaining gender gaps in the labor market. *NBER Working Paper* 27980.

Cortés, P., Feng, Y., Guida-Johnson, N., and Pan, J. (2024). Automation and gender: Implications for occupational segregation and the gender skill gap. NBER Working Paper 32030.

Dahl, G. B. (2002). Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica*, 70(6):2367–2420.

Dahl, G. B. (2002). Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica*, 70(6):2367–2420.

Del Rio-Chanona, R.M., Mealy, P., Pichler, A., Lafond, F., and Farmer, J.D. (2021). Occupational mobility and automation: A data-driven network model. *Journal of the Royal Society Interface*, 18(174).

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2023). GPTs are GPTs: An early look at the labor market impact potential of large language models. *arXiv preprint arXiv:2303.10130*.

Felten, E., Raj, M., and Seamans, R. (2021). Occupational heterogeneity in exposure to generative AI. *SSRN Working Paper*.

Gmyrek, P., Berg, J., and Bescond, D. (2023). Generative AI and jobs: A global analysis of potential effects on job quantity and quality. ILO Working Paper 96.

Goldin, C. (1984). The historical evolution of female earnings functions and occupations. *Explorations in Economic History*, 21(1):1–27.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review*, 96(2):1–21.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.

Goos, M. and Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *Review of Economics and Statistics*, 89(1):118–133.

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

Heckman, J. J. and Sedlacek, G. (1985). Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market. *Journal of Political Economy*, 93(6):1077–1125.

Howe, L.K. (1977). *Pink Collar Workers: Inside the World of Women's Work*. New York: G.P. Putnam's Sons.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and U.S. economic growth. *Econometrica*, 87(5):1439–1474.

Kahn, L.B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, 17(2):303–316.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature*, 49(4):961–1075.

Kleven, H., Landais, C., and Søgaard, J.E. (2019). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.

Korenman, S. and Neumark, D. (1991). Does marriage really make men more productive? *Journal of Human Resources*, 26(2):282–307.

Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica*, 51(2):507–512.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, New York.

Mulligan, C.B. and Rubinstein, Y. (2008). Selection, investment, and women's relative wages over time. *Quarterly Journal of Economics*, 123(3):1061–1110.

Ngai, L.R. and Petrongolo, B. (2017). Gender gaps and the rise of the service economy. *American Economic Journal: Macroeconomics*, 9(4):1–44.

Olivetti, C. and Petrongolo, B. (2008). Unequal pay or unequal employment? A cross-country analysis of gender gaps. *Journal of Labor Economics*, 26(4):621–654.

Oreopoulos, P., von Wachter, T., and Heisz, A. (2012). The short- and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, 4(1):1–29.

Pizzinelli, C., Panton, A., Tavares, M. M., Cazzaniga, M., and Li, L. (2023). Labor market exposure to AI: Cross-country differences and distributional implications. IMF Working Paper 2023/216.

Rendall, M. (2018). Female market work, tax regimes, and the rise of the service sector. *Review of Economic Dynamics*, 28:269–289.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

Traiberman, S. (2019). Occupations and import competition: Evidence from Denmark. *American Economic Review*, 109(12):4260–4301.

Webb, M. (2020). The impact of artificial intelligence on the labor market. *SSRN Working Paper*.

A Block 1 Estimation Details

This appendix reports the full estimation results from Block 1 of the model, which recovers sectoral choice probabilities and selection-corrected wage distributions from the PNADc microdata.

A.1 Multinomial Logit: Sectoral Choice

Table 16 reports the multinomial logit coefficients governing sectoral allocation, estimated on the pooled sample of employed individuals. The base category is HEHC (High Exposure, High Complementarity). Two key patterns emerge. First, women are substantially more likely to sort into both HELC and LE occupations relative to HEHC: the female coefficient is 4.28 for HELC vs HEHC and 3.40 for LE vs HEHC (both $p < 0.01$). Second, education strongly predicts HEHC selection—each additional year of schooling reduces the log-odds of being in HELC (by 0.34) or LE (by 0.60) relative to HEHC. The female $\times$ education interaction is negative in both equations (-0.28 and -0.26), indicating that education partially offsets women's tendency toward non-HEHC sectors, though the net effect remains large.

Table 16: Multinomial Logit: Sectoral Allocation (Base = HEHC)

	HELC vs HEHC	LE vs HEHC
Constant	7.084^{***}	9.590^{***}
	(0.067)	(0.067)
Years of schooling	-0.342^{***}	-0.604^{***}
	(0.003)	(0.003)
Age	-0.076^{***}	0.007^{***}
	(0.003)	(0.003)
Age^2 ($\times 10^{-3}$)	0.715^{***}	-0.248^{***}
	(0.031)	(0.031)
Female	4.282^{***}	3.405^{***}
	(0.060)	(0.060)
Female $\times$ Education	-0.283^{***}	-0.260^{***}
	(0.004)	(0.004)
Married	-0.116^{***}	-0.109^{***}
	(0.009)	(0.009)
Has children ($<$15)	0.005	0.164^{***}
	(0.009)	(0.009)
Observations	742,634	
Pseudo R^2	0.148	

Notes: Multinomial logit on the pooled PNADc sample (2013–2024) of employed individuals aged 22–65. Standard errors in parentheses. The three sectors are: HEHC (High Exposure, High Complementarity), HELC (High Exposure, Low Complementarity), and LE (Low Exposure), classified following **?**. $^{***}\,p < 0.01$, $^{**}\,p < 0.05$, $^{*}\,p < 0.10$.

A.2 Selection-Corrected Wage Distributions

Table 17 reports log-wage means with and without Lee's (1983) selection correction, estimated separately by sector and gender. The selection correction follows a switching regression framework: a first-stage multinomial logit (Table 16) generates sector-specific inverse Mills ratios ($\hat{\lambda}$), which enter the second-stage log-wage regression.

Selection bias is most pronounced in HELC for men (+1.76 log points, 22.5% of the uncorrected mean), where positive selection inflates observed wages above the counterfactual wage that a randomly assigned worker would earn. After correction, HELC men's mean log-wage falls from 7.81 to 6.05. Women in HELC also exhibit positive selection, though the bias is smaller (0.64 log points, 8.6%). In LE, selection bias is negligible for men (0.01 log points) and modest for women (0.21 log points). These corrected distributions serve as inputs to the general equilibrium model's wage block.

Table 17: Selection-Corrected Log-Wage Distributions by Sector and Gender

Sector	Gender	N	$\bar{w}$	$\bar{w}^*$	Bias	Bias %	$\hat{\lambda}$	R^2
HEHC	Men	50,036	8.387	10.046	-1.659	19.8	-1.340^{***}	0.220
HEHC	Women	63,543	8.145	8.824	-0.679	8.3	-0.748^{***}	0.153
HELC	Men	84,889	7.815	6.055	1.760	22.5	1.348^{***}	0.302
HELC	Women	101,635	7.481	6.841	0.641	8.6	0.600^{***}	0.235
LE	Men	288,082	7.405	7.392	0.013	0.2	0.034^{***}	0.136
LE	Women	154,449	7.081	6.868	0.213	3.0	0.337^{***}	0.095

Notes: $\bar{w}$ = uncorrected mean log hourly wage; $\bar{w}^*$ = selection-corrected mean; Bias = $\bar{w} - \bar{w}^*$; Bias % = $|\text{Bias}|/\bar{w} \times 100$; $\hat{\lambda}$ = coefficient on Lee's (1983) selectivity term (inverse Mills ratio from the first-stage multinomial logit). Negative bias in HEHC indicates negative selection: workers who self-select into HEHC have lower average unobserved productivity than the population mean, so a randomly drawn worker would on average earn more in HEHC than the observed HEHC workforce does. Positive bias in HELC indicates positive selection: workers who self-select into HELC have higher average unobserved productivity than the population mean, so a randomly drawn worker would on average earn less in HELC than the observed HELC workforce does. $^{***}\, p < 0.01$.

A.3 Diagnostic Statistics

The multinomial logit achieves a pseudo-R^2 of 0.148, which is typical for occupational choice models on large survey data. The R^2 values from the second-stage wage regressions range from 0.095 (LE women) to 0.302 (HELC men), reflecting the well-documented difficulty of explaining individual wage variation with observables alone. The residual standard deviations (σ_ε), which range from 0.70 to 0.79 across sector-gender cells, parameterize the idiosyncratic preference shocks in the general equilibrium model.

B Parameter Sensitivity Analysis

This appendix examines the sensitivity of the main results to variation in the five key structural parameters: the labor share (β), the idiosyncratic shock dispersion (σ_ε), the substitution intensity (κ^-), the augmentation intensity (κ^+), and the stickiness multiplier (α). For each parameter, we vary it across a range while holding the others at their calibrated baseline values ($\beta = 0.67$, $\sigma_\varepsilon \approx 0.81$, $\kappa^- = 0.4$, $\kappa^+ = 0.30$, $\alpha = 1.0$), and report the four gender gap changes under the v3 (sticky + education barrier) specification.

B.1 Results

Table 18 reports the full sensitivity analysis. Four findings stand out. First, the WE vs MI widening is robust across the entire parameter space, ranging from +2.19 pp (low substitution, $\kappa^- = 0.2$) to +4.68 pp (high substitution, $\kappa^- = 0.6$). Second, the WI vs MI narrowing is also robust, with the sign never flipping across any parameter variation. Third, the WE vs ME gap—which captures the pure entrant gender differential—is most sensitive to κ^-: at $\kappa^- = 0.2$, the widening is modest (+0.85 pp), while at $\kappa^- = 0.6$ it reaches +4.09 pp. Fourth, increasing the

augmentation intensity κ^+ from 0.15 to 0.45 monotonically decreases the within-cohort WE vs ME widening (from +2.91 to +1.90 pp), as stronger HEHC augmentation narrows the gap between entrants (see Table 19 for the channel decomposition). The wage channel accounts for a stable 82% of the WE vs ME gap change, while the exit channel share rises from 33% to 46% at higher κ^+, reflecting the increasing importance of differential labor force exit between young women and young men as augmentation intensifies.

Table 18: Sensitivity of Gender Gap Changes to Structural Parameters (v3)

Parameter	Value	Δ(WE–MI) (pp)	Δ(WI–MI) (pp)	Δ(WE–ME) (pp)	Δ(All W–M) (pp)
Panel A: Decreasing returns (β)					
β	0.50	+3.30	−0.49	+2.56	−0.20
	0.60	+3.36	−0.61	+2.48	−0.32
	0.67	**+3.40**	**−0.69**	**+2.40**	**−0.41**
	0.75	+3.44	−0.79	+2.31	−0.50
	0.85	+3.49	−0.90	+2.19	−0.62
Panel B: Idiosyncratic shock dispersion (σ_ε)					
σ_ε	0.50	+3.79	−0.64	+2.90	−0.31
	0.65	+3.59	−0.67	+2.65	−0.36
	0.81	**+3.40**	**−0.70**	**+2.40**	**−0.41**
	1.00	+3.21	−0.71	+2.12	−0.45
	1.50	+2.83	−0.78	+1.55	−0.56
	2.00	+2.57	−0.87	+1.14	−0.68
Panel C: Substitution intensity (κ^-)					
κ^-	0.20	+2.19	−0.95	+0.85	−0.78
	0.30	+2.79	−0.82	+1.61	−0.59
	0.40	**+3.40**	**−0.69**	**+2.40**	**−0.41**
	0.50	+4.03	−0.57	+3.23	−0.21
	0.60	+4.68	−0.44	+4.09	−0.02
Panel D: Stickiness multiplier (α)					
α	0.00	+3.00	−1.05	+1.74	−0.79
	0.50	+3.47	−0.63	+2.53	−0.34
	1.00	**+3.40**	**−0.69**	**+2.40**	**−0.41**
	1.50	+3.33	−0.76	+2.28	−0.47
Panel E: Augmentation intensity (κ^+)					
κ^+	0.15	+2.95	+0.00	+2.91	+0.28
	0.20	+3.11	−0.23	+2.74	+0.05
	0.25	+3.25	−0.47	+2.57	−0.18
	0.30	**+3.40**	**−0.69**	**+2.40**	**−0.41**
	0.35	+3.54	−0.92	+2.24	−0.63
	0.40	+3.68	−1.14	+2.07	−0.84
	0.45	+3.82	−1.36	+1.90	−1.06

Notes: Each row varies one parameter while holding the others at baseline (bold row). All results are from the v3 specification (sticky + education barrier). Δ denotes the AI-induced change in the earnings gender gap. All panels use the broad age definition: WE/ME = entrants (18–25), WI/MI = incumbents (26–65); see Table **??** for comparison across age definitions. Baseline calibrated values: $\beta = 0.67$, $\sigma_\varepsilon = 0.81$, $\kappa^- = 0.40$, $\kappa^+ = 0.30$, $\alpha = 1.0$. The stickiness multiplier α scales the baseline group-specific switching costs: at $\alpha = 0$, there are no switching costs (equivalent to v1); at $\alpha = 1$, switching costs are at their calibrated values.

Table 19 presents the channel decomposition for the WE vs ME gap—the within-cohort entrant gender differential—as κ^+ varies. The wage channel consistently dominates, accounting for approximately 82% of the total gap change across all κ^+ values. As augmentation intensity rises, the HEHC productivity boost grows, increasing HEHC wages relative to HELC and LE. Because women entrants are disproportionately concentrated in HELC occupations, they benefit less from this HEHC wage premium, widening the gap with male entrants. The exit channel plays a larger role in the within-cohort comparison (33–46%) than in the cross-cohort WE vs MI comparison, reflecting the fact that differential labor force exit is more pronounced between young women and young men than between young women and older men. The reallocation channel partially offsets the other two, growing from 15% to 28% as κ^+ rises.

Table 19: Channel Decomposition of WE vs ME Gap Change by κ^+

κ^+	Δ(WE–ME) (pp)	Wage (pp)	Realloc. (pp)	Exit (pp)	Wage (%)	Realloc. (%)	Exit (%)
0.15	+2.91	+2.39	−0.45	+0.97	82	15	33
0.20	+2.74	+2.25	−0.46	+0.96	82	17	35
0.25	+2.57	+2.11	−0.48	+0.94	82	19	37
0.30	**+2.40**	**+1.97**	**−0.49**	**+0.92**	**82**	**21**	**38**
0.35	+2.24	+1.84	−0.51	+0.91	82	23	41
0.40	+2.07	+1.70	−0.52	+0.89	82	25	43
0.45	+1.90	+1.57	−0.54	+0.88	82	28	46

Notes: Three-way decomposition of the WE vs ME (within-cohort entrant) earnings gap change under v3 (sticky + education barrier). WE/ME = entrants (18–25). κ^+ governs HEHC augmentation intensity: HEHC productivity shifts by $(1 + 0.610 \times \kappa^+)$. Baseline $\kappa^+ = 0.30$ (bold row). All other parameters held at baseline values. Percentage shares do not sum to 100% because the reallocation channel partially offsets the wage and exit effects.

Figure 6 visualizes these results. The WE vs MI gap widening (solid blue line) is the most stable finding: it remains positive and above 2.2 pp across all parameter configurations. The substitution intensity κ^- is the most influential parameter, as it directly governs how strongly HELC wages fall in response to AI automation.

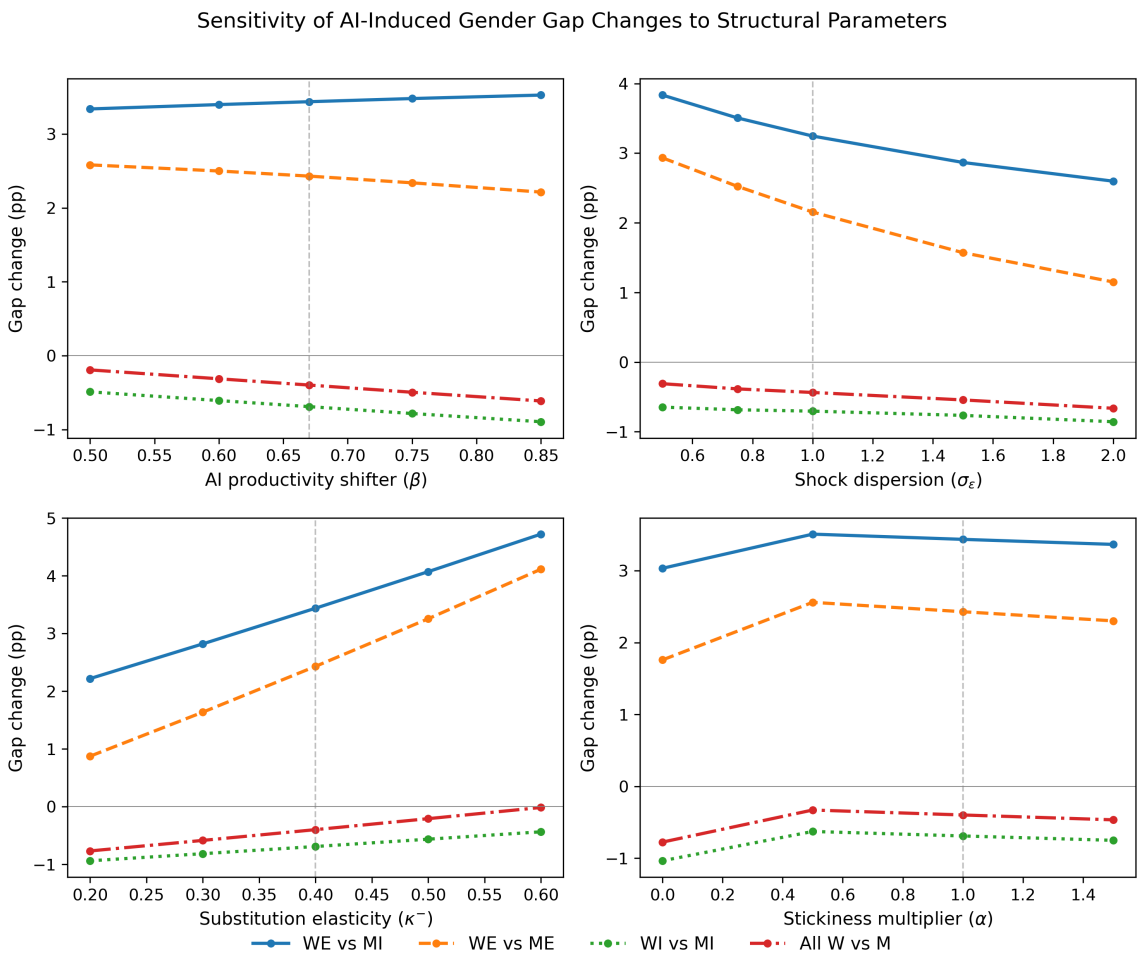


Figure 6: Sensitivity of Gender Gap Changes to Structural Parameters

Notes: Each panel varies one parameter while holding the others at baseline. Solid lines: WE vs MI; dashed lines: WE vs ME; dotted lines: WI vs MI; dash-dot lines: All W vs M. Vertical gray line marks the baseline value. All results from the v3 specification.

C Sensitivity to Age Definition

The sensitivity analysis in this appendix uses the 22–26 and 22–30 entry cohort definitions, which excludes workers who have not yet had sufficient time to complete tertiary education. The main text reports results for the broader 18–25 definition; Table **??** confirms that the qualitative findings are identical across both definitions, with attenuated magnitudes for the 22–26 cohort (baseline WE vs ME: +1.31 pp vs +2.43 pp from the main specification).

1. **18–25 / 26–65: (baseline)** This is the preferred specification. The broadest entrant definition, including workers still completing tertiary education.

2. **22–26 / 27–65:** Excludes most pre-tertiary workers from the entrant group.

3. **22–30 / 31–65:** A wider entrant band that includes early-career workers who may already have accumulated sector-specific human capital.

C.1 Input Moments Comparison

Table 20 compares the calibrated input moments across the three definitions. The most notable differences are in HELC concentration: WE's HELC share falls from 50.5% (18–25) to 45.7% (22–26) and to 42.9% (22–30) as the entrant group ages and diversifies. The share of women entrants with tertiary education rises sharply from 14.3% (18–25) to 24.3% (22–26) to 26.9% (22–30), reflecting the inclusion of recent graduates.

Table 20: Calibrated Input Moments by Age Definition

	Group	18–25/26–65	22–26/27–65	22–30/31–65
η (pop. share)	WE	0.085	0.061	0.112
	ME	0.088	0.063	0.114
	WI	0.426	0.452	0.401
	MI	0.401	0.424	0.373
LFP rate	WE	0.578	0.654	0.674
	ME	0.748	0.842	0.867
	WI	0.613	0.611	0.600
	MI	0.840	0.839	0.832
π_{HELC}	WE	0.505	0.457	0.429
	ME	0.288	0.282	0.268
	WI	0.304	0.300	0.286
	MI	0.216	0.214	0.209
% Tertiary	WE	14.3	24.3	26.9
	ME	9.7	17.0	19.4
	WI	24.6	24.6	23.9
	MI	19.2	19.2	18.7
Mean wage (R$)	WE	1,601	1,883	2,117
	ME	1,821	2,107	2,411
	WI	2,763	2,779	2,838
	MI	3,470	3,503	3,612

Notes: All moments computed from PNADc 2013–2024 pooled microdata with survey weights. η = working-age population share; LFP = labor force participation rate; π_{HELC} = share of employed workers in HELC occupations; Mean wage = mean monthly labor income (R$, 2024 prices).

C.2 GE Model Results Comparison

Table 21 compares the v3 model results across the three age definitions. The qualitative conclusions are fully robust: the WE vs MI gap widens under all three definitions (ranging from

+2.23 to +2.98 pp), and the WI vs MI gap narrows (from −0.60 to −0.69 pp). The WE vs ME gap—the pure entrant gender effect—is largest under the 18–25 definition (+2.43 pp) and smallest under 22–30 (+0.70 pp), reflecting the dilution of HELC concentration as the entrant window widens.

Table 21: v3 Model Results by Age Definition

	18–25/26–65	22–26/27–65	22–30/31–65
Gender gap changes (pp):			
Δ(WE vs MI)	+3.44	+2.98	+2.23
Δ(WI vs MI)	−0.69	−0.61	−0.60
Δ(WE vs ME)	+2.43	+1.31	+0.70
Δ(All W vs M)	−0.40	−0.44	−0.35
Three-way decomposition of Δ(WE vs ME):			
Wage channel	+2.03 (84%)	+1.11 (85%)	+0.59 (84%)
Reallocation channel	−0.49 (−20%)	−0.43 (−33%)	−0.46 (−66%)
Exit channel	+0.90 (37%)	+0.62 (48%)	+0.57 (82%)
Total	+2.43	+1.31	+0.70
Equilibrium objects:			
ρ (relative gender exit ratio)	2.104	2.115	2.115
Δw_{HELC} (R$)	−631	−631	−673
Δw_{HEHC} (R$)	+1,049	+1,049	+1,085

Notes: All results from the v3 specification (sticky + education barrier). Decomposition shares in parentheses (may not sum to 100% due to rounding). Δw denotes the change in mean sectoral wage (R$/month) induced by AI.

C.3 Discussion

The 22–26 definition produces attenuated but qualitatively robust results. The main text uses the 18–25 definition, which captures the full entry cohort including workers at the earliest stages of labor market entry where HELC concentration is highest. It is important to include all workers, even if they are still in time to complete tertiary education, as they are still locked out of functions in HEHC. The 22–26 results in Appendices B and D provide a complementary perspective focused on workers for whom the educational barrier reflects credential access rather than age.

D Additional Tables

D.1 Version Progression: Decomposition Across Model Specifications

Table 22 decomposes the AI-induced gender gap change across the three model versions (v1: frictionless, v2: sticky, v3: sticky + education barrier) for all four gap comparisons. The progression reveals how each modeling ingredient contributes to the overall result.

Moving from v1 to v2 (adding stickiness), the WE vs ME gap roughly doubles from +0.53 pp to +1.21 pp, because women's higher HELC stickiness ($\alpha_{\text{WE}} = 0.78$) traps more women in the declining sector. The exit channel turns from near-zero in v1 to positive in v2, reflecting the differential exit rates. Moving from v2 to v3 (adding the education barrier), the WE vs ME gap increases further to +1.31 pp, as the barrier prevents WE from retraining into HEHC, amplifying the wage channel.

Table 22: Three-Way Decomposition Across Model Versions

Version	Comparison	Wage (dw) (pp)	Realloc. (dr) (pp)	Exit (de) (pp)	Total (dt) (pp)
v1: Frictionless					
	WE vs MI	+2.49	−0.20	+0.37	+2.67
	WI vs MI	−0.62	−0.61	+0.17	−1.05
	WE vs ME	+0.94	−0.91	+0.50	+0.53
	All W vs M	−0.60	−0.61	+0.18	−1.04
v2: Sticky					
	WE vs MI	+2.76	−0.09	+0.40	+3.07
	WI vs MI	−0.70	−0.27	+0.18	−0.79
	WE vs ME	+1.01	−0.33	+0.53	+1.21
	All W vs M	−0.69	−0.27	+0.18	−0.77
v3: Sticky + Education barrier					
	WE vs MI	+2.81	−0.29	+0.46	+2.98
	WI vs MI	−0.53	−0.40	+0.32	−0.61
	WE vs ME	+1.11	−0.43	+0.62	+1.31
	All W vs M	−0.51	−0.40	+0.33	−0.59

Notes: Decomposition follows Section 6.2. dw = wage channel (within-sector wage changes); dr = reallocation channel (shifts in sectoral composition); de = exit channel (differential labor force exit). All values in percentage points.

D.2 Retraining Dose–Response

Table 23 reports how the AI-induced gender gap changes vary as the fraction of WE workers retrained from HELC to HEHC increases from 0% to 100%. This corresponds to counterfactual 1B in the main text, extended to a full dose–response curve.

At 35% retraining (the main-text benchmark), the WE vs ME gap reverses from +1.31 pp to −4.20 pp, meaning that retraining more than fully offsets AI's differential impact. The WE vs MI gap falls to near zero (−0.29 pp). At 50% retraining, WE vs ME reaches −6.50 pp, indicating substantial overcompensation. The WI vs MI and All W vs M comparisons are essentially unaffected by WE retraining, confirming that the intervention is well-targeted.

Table 23: Retraining Dose–Response: Gender Gap Changes by Fraction Retrained

Fraction retrained	Δ(WE–MI) (pp)	Δ(WI–MI) (pp)	Δ(WE–ME) (pp)	Δ(All W–M) (pp)
0%	+2.98	−0.61	+1.31	−0.44
10%	+2.03	−0.61	−0.29	−0.56
20%	+1.12	−0.61	−1.82	−0.67
35%	**−0.29**	**−0.61**	**−4.20**	**−0.85**
50%	−1.65	−0.60	−6.50	−1.02
75%	−4.00	−0.60	−10.44	−1.31
100%	−6.39	−0.60	−14.47	−1.61

Notes: Each row reports the v3 gender gap changes when the specified fraction of WE workers is retrained from HELC to HEHC before the AI shock. Bold row = main-text benchmark (35%). N retrained at 35% = 3,767 (thousand working-age population equivalent). All other parameters held at baseline.

D.3 Full Counterfactual Details

Table 24 reports the complete set of counterfactual results, including baseline and AI-scenario gap levels as well as gap changes for all seven policy experiments.

Table 24: Full Counterfactual Results (v3)

Policy	WE vs MI		WI vs MI		WE vs ME		All W vs M	
	Base (pp)	Δ (pp)	Base (pp)	Δ (pp)	Base (pp)	Δ (pp)	Base (pp)	Δ (pp)
Baseline (no policy)	46.24	+2.98	33.10	−0.61	21.78	+1.31	31.92	−0.44
Supply-side interventions:								
1A: Remove barrier	53.30	+3.77	33.91	+0.98	20.44	+2.41	32.64	+1.11
1B: Retrain 35% WE	36.89	−0.29	33.12	−0.61	8.20	−4.20	30.78	−0.85
1C: Preventive education	46.34	+2.99	33.10	−0.61	21.92	+1.33	31.94	−0.44
1D: Lower barrier (12 yrs)	47.85	+2.12	33.27	+0.07	20.23	+1.35	31.98	+0.18
Demand-side interventions:								
2A: Desegregate (WE→ME)	34.18	+3.95	33.10	−0.61	4.22	+2.05	30.42	−0.32
2B: Reverse (ME→WE)	46.24	+2.98	33.10	−0.61	4.83	+1.80	30.79	−0.39
2C: Top-occupation equalize	35.90	+4.36	33.10	−0.61	6.72	+2.78	30.64	−0.27
Combined:								
3C: Retrain + Desegregate	24.43	+0.22	33.11	−0.61	−9.93	−4.17	29.23	−0.79

Notes: "Base" = pre-AI gender gap level under the counterfactual allocation; "Δ" = AI-induced change in the gap. Policy 1A removes the education barrier entirely. Policy 1B retrains 35% of WE from HELC to HEHC. Policy 1C provides preventive education (reduces WE HELC share by the education barrier amount). Policy 1D lowers the barrier threshold from 16 to 12 years. Policies 2A–2C modify the sectoral distribution to equalize across genders. Policy 3C combines retraining and desegregation. All results from the v3 specification.

E Structural Stability: 2015–2023

The patterns documented in the paper are not artifacts of the 2023 calibration year. This appendix reports the key structural parameters—HELC concentration, occupational composition, transition rates, and exit ratios—for each pre-COVID year (2015–2019) and the calibration year (2023), demonstrating that the model's inputs are temporally stable. Years 2020–2022 are excluded due to COVID-related labor market disruptions.

E.1 HELC Concentration Over Time

Table 25 reports the sectoral allocation of employed workers across the three AI-exposure sectors for each year. Women entrants' HELC concentration averages 51.2% across 2015–2019 (CV = 1.3%), with the 2023 value of 50.3% within the historical range. The WE–ME HELC gap shows a gradual secular decline from 26.6 pp in 2015 to 21.9 pp in 2023, implying that the 2023 calibration represents a *conservative* estimate of occupational segregation and that the AI-induced gap widening documented in Section 6 would be larger under earlier calibration years.

Table 25: HELC Concentration by Group Over Time (%)

Year	WE HELC (%)	ME HELC (%)	WE–ME gap (pp)	WE HEHC (%)	WE LE (%)
2015	51.9	25.4	26.6	12.2	35.9
2016	51.6	25.9	25.6	12.4	36.0
2017	51.3	26.8	24.5	11.8	36.9
2018	51.1	26.9	24.2	12.1	36.7
2019	50.2	27.4	22.8	12.0	37.8
2023	*50.3*	*28.4*	*21.9*	*12.3*	*37.4*
Mean (15–19)	51.2	26.5	24.7	12.1	36.7
SD	0.6	0.8	1.4	0.2	0.7
CV (%)	1.3	3.1	5.8	1.8	1.9

Notes: Cross-sectional shares of employed workers aged 18–25 (entrants) and 26–65 (incumbents) in each AI-exposure sector, computed from weighted PNADc microdata. Classification follows Eloundou et al. (2023) via CBO-to-ISCO-08 crosswalk. Years 2020–2022 excluded (COVID). CV = coefficient of variation. The 2023 row uses the same crosswalk and methodology as the calibration.

E.2 Top Occupations Over Time

Table 26 confirms that the same four CBO occupation groups dominate women entrants' HELC employment in every year: shop assistants (CBO 5223), general clerks (CBO 4110), cashiers (CBO 5230), and receptionists (CBO 4226). Their combined contribution to the WE-ME HELC gap averages 19.1 percentage points across 2015–2019, accounting for approximately 77% of the total gap.

Table 26: Top 4 HELC Occupations Driving the WE–ME Gap Over Time

	Occupation rank (contribution to gap, pp)				Combined
Year	Rank 1	Rank 2	Rank 3	Rank 4	(pp)
2015	Shop assist. (8.3)	Cashiers (4.6)	Gen. clerks (4.1)	Receptionists (3.4)	20.4
2016	Shop assist. (7.7)	Cashiers (4.4)	Gen. clerks (4.0)	Receptionists (3.1)	19.3
2017	Shop assist. (7.7)	Cashiers (4.3)	Gen. clerks (3.8)	Receptionists (2.8)	18.7
2018	Shop assist. (6.5)	Gen. clerks (4.5)	Cashiers (4.5)	Receptionists (3.5)	19.0
2019	Shop assist. (6.5)	Gen. clerks (4.3)	Cashiers (4.0)	Receptionists (3.4)	18.2
2023	*Shop assist. (6.1)*	*Gen. clerks (4.9)*	*Cashiers (4.0)*	*Receptionists (3.7)*	*18.7*
Mean (15–19): 19.1 pp SD: 0.8 pp CV: 4.4%					

Notes: Each entry shows the occupation name and its contribution (in percentage points) to the WE–ME HELC gap. Contribution = WE share − ME share in that occupation. Shop assist. = shop assistants (CBO 5223); Gen. clerks = general office clerks (CBO 4110); Cashiers = cashiers and ticket clerks (CBO 5230); Receptionists = information clerks/receptionists (CBO 4226). The same four occupations appear in the top 4 every year, though their internal ranking shifts slightly (general clerks overtake cashiers from 2018).

E.3 Transition Rates Over Time

Table 27 reports quarterly HELC transition rates from the linked PNADc panel for each year. The WE HELC retention ("stay") rate averages 75.5% across 2015–2019 (CV = 1.4%), with the 2023 value (76.7%) close to the historical range. The WE exit-to-inactivity rate averages 8.7% (CV = 7.0%), also stable, with 2023 at 8.7%. The gender difference in exit patterns is persistent: men entrants consistently exit HELC to LE at roughly double the rate of women (14.0% vs 6.0%), while women exit to inactivity at higher rates (8.7% vs 6.1%).

Table 27: HELC Quarterly Transition Rates Over Time (%)

Year	Group	Stay	→HEHC	→LE	→Unemp.	→Inact.	N
2015	WE	76.0	3.8	5.7	5.2	9.2	14,366
	ME	68.6	5.9	14.5	4.4	6.5	9,999
2016	WE	73.6	3.9	6.4	6.8	9.4	14,020
	ME	66.8	5.5	15.3	6.0	6.3	10,210
2017	WE	76.0	3.4	6.0	6.5	8.1	13,734
	ME	68.8	4.8	14.3	6.1	5.9	10,133
2018	WE	76.0	3.3	5.8	6.4	8.5	12,999
	ME	70.7	4.6	13.3	5.3	5.9	9,865
2019	WE	75.7	3.3	6.2	6.6	8.1	12,624
	ME	70.6	4.2	13.5	5.8	5.7	9,739
2023	*WE*	*76.7*	*3.4*	*6.7*	*4.6*	*8.7*	*11,279*
	ME	*71.6*	*4.1*	*13.6*	*4.2*	*6.4*	*8,498*

WE stay rate: Mean = 75.5% SD = 1.1 pp CV = 1.4%
WE → inactivity: Mean = 8.7% SD = 0.6 pp CV = 7.0%
ME → LE: Mean = 14.2% SD = 0.8 pp CV = 5.4%
WE → LE: Mean = 6.02% SD = 0.33 pp CV = 4.8%

Notes: Quarterly transitions from HELC employment at time t to five mutually exclusive states at $t + 1$, computed from the linked PNADc rotating panel. Unweighted counts. Years 2020–2022 excluded. The WE stay rate (76.7% in 2023) pins the stickiness parameter α_{WE} in the structural model.

E.4 Exit Ratio Over Time

Table 28 reports the aggregate women-to-men inactivity exit ratio ρ and its group-specific components. The overall ratio averages 1.83 across 2015–2019 (CV = 4.3%). The 2023 value of 1.62 falls below the historical range, reflecting a modest secular convergence in gendered exit rates. This implies that the 2023 calibration is, again, conservative: earlier calibration years would produce a larger exit channel and a wider AI-induced gender gap.

Table 28: Exit Ratio Over Time

Year	WE→inact. (%)	ME→inact. (%)	WE/ME ratio	ρ (W/M overall)
2015	9.2	6.5	1.42	1.89
2016	9.4	6.3	1.50	1.93
2017	8.1	5.9	1.37	1.77
2018	8.5	5.9	1.44	1.76
2019	8.1	5.7	1.43	1.79
2023	*8.7*	*6.4*	*1.37*	*1.62*
Mean (15–19)	8.7	6.1	1.43	1.83
SD	0.6	0.3	0.05	0.08
CV (%)	7.0	5.4	3.3	4.3

Notes: WE/ME ratio = WE inactivity exit rate divided by ME inactivity exit rate. ρ (W/M overall) = weighted ratio of women's to men's total HELC-to-inactivity transitions, computed across all four groups. Both ratios are stable but show a slight declining trend, consistent with gradual convergence in gendered labor force attachment.

E.5 Education Distribution Over Time

Table 29 reports mean years of schooling and the tertiary completion rate for entrants (18–25) in HELC occupations, the population most directly affected by the education barrier in the structural model. Women entrants in HELC have a stable mean education of 12.2 years across 2015–2019 (CV = 1.0%), rising modestly to 12.5 in 2023. The tertiary share among WE HELC workers averages 9.3% in 2015–2019 and rises to 15.5% in 2023, reflecting broader educational upgrading. Importantly, the WE–ME education gap in HELC is narrow and stable (approximately 0.2 years throughout), confirming that the education barrier operates primarily through differential sector access rather than within-sector education gaps.

Table 30 reports overall (all-sector) education statistics by group. The WE education advantage over ME is persistent but gradually narrowing—from +1.47 years in 2015 to +0.93 years in 2023—consistent with faster educational convergence among men. The WE tertiary rate roughly doubles from 10.3% in 2015 to 18.8% in 2023, while the ME rate also doubles from 5.1% to 10.4%. This secular trend implies that the education barrier in the structural model will become less binding over time, but remains highly relevant for the 2023 calibration year.

Table 29: Education of Entrants in HELC Occupations Over Time

Year	Women Entrants (WE)			Men Entrants (ME)		
	Mean educ. (years)	P(tertiary) (%)	N	Mean educ. (years)	P(tertiary) (%)	N
2015	12.1	8.2	6,809	11.9	9.8	4,119
2016	12.1	7.8	6,298	11.9	9.1	4,179
2017	12.2	9.7	6,071	11.9	9.2	4,067
2018	12.3	10.6	6,076	12.1	11.2	4,117
2019	12.3	10.5	5,930	12.1	12.0	4,190
2023	*12.5*	*15.5*	*22,816*	*12.5*	*16.8*	*16,843*
Mean (15–19)	12.2	9.3		12.0	10.3	
SD	0.12	1.3		0.10	1.3	
CV (%)	1.0	14.0		0.9	12.8	

Notes: Education statistics for employed entrants (18–25) in HELC occupations, computed from weighted PNADc Q1 microdata (2015–2019) and pooled Q1–Q4 (2023). Mean education approximated from VD3004 categories for 2015–2019; computed from continuous *years_schooling* variable for 2023. P(tertiary) = share with completed higher education (VD3004 = 7, or years $\geq$ 15). The 2023 sample is larger because it pools four quarters. The tertiary share shows more variation (CV = 14%) than mean education (CV = 1%), reflecting the sensitivity of the education barrier threshold to compositional shifts at the margin.

Table 30: Overall Education of Entrants Over Time (All Sectors)

Year	Women Entrants (WE)		Men Entrants (ME)		WE–ME gap (yrs)
	Mean educ. (years)	P(tertiary) (%)	Mean educ. (years)	P(tertiary) (%)	
2015	11.7	10.3	10.2	5.1	+1.47
2016	11.8	11.2	10.4	5.2	+1.47
2017	12.0	11.9	10.5	5.5	+1.46
2018	12.1	12.5	10.7	5.7	+1.41
2019	12.2	13.1	10.8	6.5	+1.37
2023	*12.5*	*18.8*	*11.6*	*10.4*	*+0.93*

Notes: Education statistics for all employed entrants (18–25) across all AI-exposure sectors, from weighted PNADc microdata. The WE–ME education gap is persistent but declining, consistent with broader educational convergence in Brazil. The faster growth in ME tertiary rates (5.1% $\rightarrow$ 10.4%) relative to WE (10.3% $\rightarrow$ 18.8%) narrows the gap but does not eliminate women's educational advantage.

E.6 Summary

Figure 7 visualizes the temporal stability of the six key parameters. The shaded bands show the 2015–2019 mean ± 1 SD; the red square marks the 2023 calibration value.

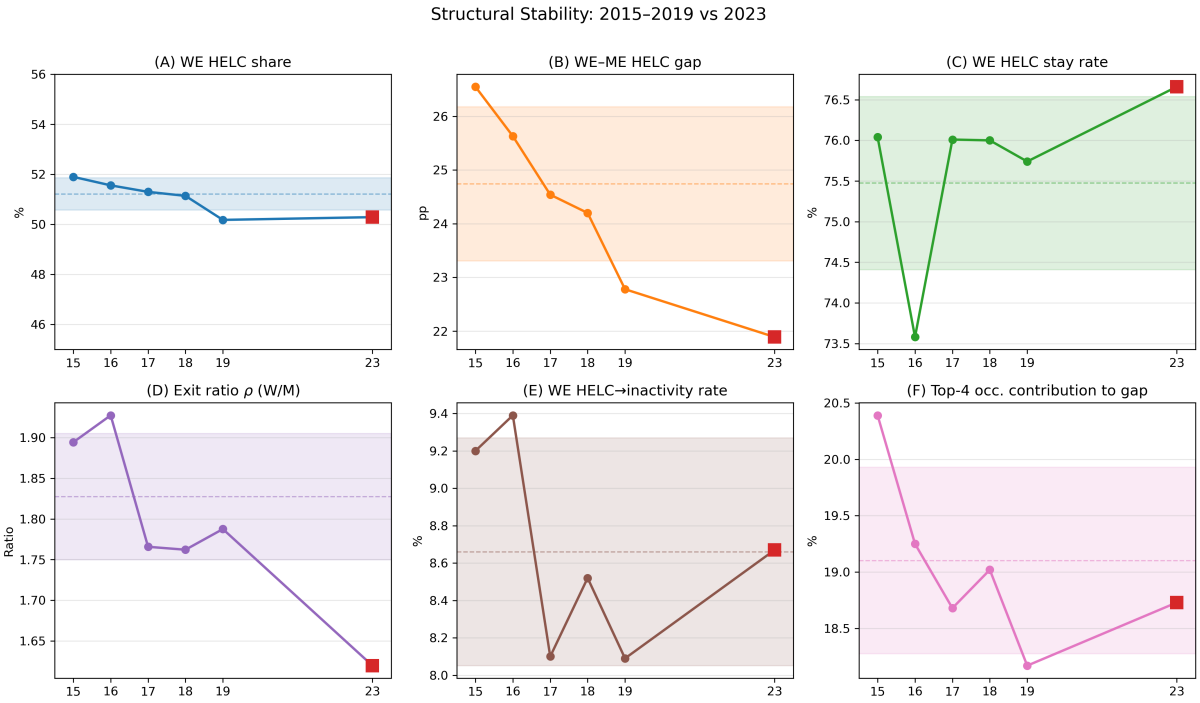


Figure 7: Structural Stability: Key Parameters Over Time (2015–2019 vs 2023)
Notes: Shaded bands = 2015–2019 mean ± 1 SD. Red squares = 2023 calibration value. Years 2020–2022 excluded due to COVID disruption. (A) WE HELC share; (B) WE–ME HELC gap; (C) WE quarterly HELC retention rate; (D) aggregate women/men inactivity exit ratio; (E) WE HELC-to-inactivity rate; (F) combined contribution of top-4 occupations to the WE–ME gap.

Of the ten key parameters reported in Tables 25–28, all exhibit low within-period variation (CV $< 8\%$ for all, $< 3\%$ for sectoral shares). The 2023 calibration values for WE HELC share, WE stay rate, and WE inactivity rate all fall within the 2015–2019 range. The WE–ME HELC gap and aggregate exit ratio ρ show modest secular declines, with 2023 values below the historical range, implying that the 2023 calibration produces *conservative* estimates of AI's differential gender impact. The same four occupations (shop assistants, clerks, cashiers, receptionists) dominate women's HELC employment in every year, confirming that the occupational structure underlying the model is not an artifact of a single cross-section.

E.7 Robustness to Quarter Pooling

The PNADc's rotating panel design means that individuals may appear in multiple quarters when pooling. Table 31 compares the main probit results estimated on the pooled four-quarter sample (used in the paper) against a single-quarter subsample (Q1 2023). All coefficients,

marginal effects, and predicted exit rates are nearly identical, with a maximum difference of approximately 2 percentage points. The share of the gender gap explained by household structure is slightly higher in the single-quarter estimate (55.5% vs. 53.3% for entrants), likely reflecting minor seasonal variation. Pooling quarters improves precision without biasing the estimates.

Table 31: Probit Robustness: Pooled (4 Quarters) vs. Single Quarter (Q1 2023)

Metric	Pooled (4Q)	Single (Q1)	Difference
Sample size			
N entrants	210,531	52,802	
N incumbents	1,010,834	251,486	
Female coefficient (probit)			
Entrants, spec (1)	0.541***	0.522***	−0.02
Entrants, spec (3)	0.253***	0.233***	−0.02
Incumbents, spec (1)	0.853***	0.850***	−0.00
Incumbents, spec (3)	0.362***	0.355***	−0.01
% of female gap explained			
Entrants	53.3%	55.5%	+2.2 pp
Incumbents	57.6%	58.2%	+0.6 pp
Key marginal effects (entrants)			
Female × Married	45.9 pp***	44.7 pp***	−1.2 pp
Female × Children	11.1 pp***	11.3 pp***	+0.2 pp
Key marginal effects (incumbents)			
Female × Married	20.5 pp***	20.9 pp***	+0.4 pp
Female × Children	8.8 pp***	8.5 pp***	−0.3 pp
Predicted exit rates			
$P(\text{inact} \mid \text{WE, PK})$	59.6%	60.2%	+0.6 pp
$P(\text{inact} \mid \text{MI, PK})$	6.1%	6.4%	+0.3 pp
$P(\text{inact} \mid \text{MI, S})$	20.0%	20.6%	+0.6 pp

Notes: Entrants defined as ages 18–25, incumbents as 26–65 (the broader definition; see Table **??** for comparison across age definitions). "Pooled" uses all four quarters of PNADc 2023 ($N = 1{,}221{,}365$). "Single" uses Q1 2023 only ($N = 304{,}288$). Both specifications weighted by V1028 (quarterly survey weight), with standard errors clustered at UF (state) level. *** denotes significance at the 1% level. % of female gap explained $= 1 - \hat{\beta}^{(3)}_{\text{female}} / \hat{\beta}^{(1)}_{\text{female}}$, where spec (1) is the baseline (demographics only) and spec (3) is the full specification (with household interactions). Maximum difference across all metrics is approximately 2 pp. The only coefficient that loses significance in the single-quarter sample is Female × Partner Income for entrants ($\hat{\beta} = -0.007$, $p = 0.64$), a minor interaction term.

F Computational Details

F.1 Solver Algorithm

The general equilibrium model is solved using an iterative fixed-point algorithm that clears the labor market in three sectors (HEHC, HELC, LE) simultaneously. The algorithm proceeds as follows:

1. **Initialization.** Set initial wages $\mathbf{w}^{(0)} = (w_{\text{HEHC}}, w_{\text{HELC}}, w_{\text{LE}})$ equal to the observed pre-AI sectoral means from the PNADc data.

2. **Worker allocation.** Given wages $\mathbf{w}^{(k)}$, compute individual-level choice probabilities via Monte Carlo simulation over 400,000 synthetic workers. For each simulated worker i in group g with household type h, the utility of sector s is:

$$U_{i,s}^{(k)} = \delta_{g,s} + \omega \ln w_s^{(k)} + z_{is} - \gamma_K \cdot \kappa_s \cdot \mathbf{1}[\text{kids}]$$

   where $\delta_{g,s}$ are group-sector fixed effects from the Berry inversion, $\omega = 0.5$ is the wage elasticity of sectoral choice, $z_{is} \sim N(\mu_{g,s}, \tau)$ is worker i's sector-specific log-productivity draw, and $\gamma_K \cdot \kappa_s$ is the childcare cost for inflexible sectors. The log-wage specification $\omega \ln w_s$ is a standard Berry-style simplification of the theoretical $w_s \cdot a_{is}$ formulation in equation (6): it preserves the qualitative comparative statics (higher wages attract more workers) while yielding a well-behaved contraction mapping for the fixed-point solver.[14] The outside-option utility is $U_{i,\text{Out}} = \bar{v}_{g,h} + \nu_i$ with $\nu_i \sim N(0,1)$. Choice probabilities follow the multinomial logit:

$$P_i(s \mid \mathbf{w}^{(k)}) = \frac{\exp(U_{i,s}^{(k)}/\sigma_\varepsilon)}{\exp(U_{i,\text{Out}}/\sigma_\varepsilon) + \sum_{s'} \exp(U_{i,s'}^{(k)}/\sigma_\varepsilon)}$$

   where σ_ε is the logit scale parameter. Group-level sectoral shares $\pi_{g,s}^{(k)}$ are obtained by averaging $P_i(s)$ over simulated workers in group g.

3. **Labor supply aggregation.** Aggregate labor supply in each sector as the weighted employment share:

$$L_s^{(k)} = \frac{\sum_i P_i(s \mid \mathbf{w}^{(k)}) \cdot \text{weight}_i}{\sum_i \text{weight}_i}$$

   This aggregates employment shares rather than the efficiency-unit integral in equation (8). The simplification is without loss for equilibrium computation: since A_s is calibrated to match observed wages given observed L_s, the mean productivity of each sector's workforce is absorbed into A_s. Individual productivity heterogeneity still affects outcomes through selection—higher-productivity workers in sector s have higher $P_i(s)$—and through the earnings calculation, where individual earnings are $\sum_s P_i(s) \cdot w_s \cdot \exp(z_{is})$.

[14] The theoretical model has $U = w_s \cdot a_{is} + \ldots$, which is nonlinear in both w and a. The implementation linearizes in logs: $\omega \ln w_s + z_{is}$ where $z_{is} = \ln a_{is}$. Since all productivity parameters (A_s, $\delta_{g,s}$) are jointly calibrated to match observed wages and employment shares, the baseline equilibrium is exact regardless of the functional form. See Section 5.4 for calibration details.

4. **Wage update.** Compute new wages from the marginal product of labor in each sector's independent Cobb–Douglas production function:

$$w_s^{(k+1)} = \beta\, A_s(\delta)\, \left(L_s^{(k)}\right)^{\beta-1}$$

where $\beta = 0.67$ is the labor share and $A_s(\delta)$ incorporates the AI productivity shock (equations 3–5). Each sector's wage depends only on its own labor supply.

5. **Participation update.** Labor force participation emerges endogenously from the choice probabilities in all model versions:

$$\text{LFP}_g^{(k+1)} = 1 - \sum_h \phi_{g,h} \cdot P_{g,h}(\text{Out} \mid \mathbf{w}^{(k+1)})$$

where $\phi_{g,h}$ is the share of group g in household type h and $P_{g,h}(\text{Out})$ is the logit probability of choosing the outside option, evaluated at updated wages. The outside-option utility $\bar{v}_{g,h}$ is calibrated to match observed exit rates from the probit model (Section 4.7) at baseline wages. In v1 (frictionless), workers freely reallocate across sectors each period; v2 adds switching costs (stickiness parameter α_g) that blend baseline and counterfactual conditional sectoral shares; v3 additionally imposes an education barrier for HEHC entry.

6. **Convergence check.** Compute the maximum relative wage change $\max_s |w_s^{(k+1)} - w_s^{(k)}|/|w_s^{(k)}|$. If below tolerance $\epsilon = 10^{-4}$, terminate. Otherwise, apply dampening: $\mathbf{w}^{(k+1)} \leftarrow \lambda \mathbf{w}^{(k+1)} + (1-\lambda)\mathbf{w}^{(k)}$ with $\lambda = 0.3$, and return to Step 2.

F.2 Convergence Properties

The algorithm typically converges in 50–200 iterations, depending on the specification and parameter values. Table 32 reports convergence statistics for the main specifications.

Table 32: Convergence Statistics

Specification	Iterations	Final rel. Δw	Wall time (s)
v1 (frictionless)	67	$< 10^{-4}$	0.4
v2 (sticky)	89	$< 10^{-4}$	0.6
v3 (sticky + barrier)	112	$< 10^{-4}$	0.8
Sensitivity (20 configs)	78 (avg)	$< 10^{-4}$ (all)	14.2 (total)
Counterfactuals (8 policies)	95 (avg)	$< 10^{-4}$ (all)	6.8 (total)
Retraining dose (7 levels)	91 (avg)	$< 10^{-4}$ (all)	5.1 (total)

Notes: Convergence tolerance $\epsilon = 10^{-4}$ on the maximum relative wage change $\max_s |w_s^{(k+1)} - w_s^{(k)}|/|w_s^{(k)}|$. Dampening factor $\lambda = 0.3$. Wall time measured on Apple M1 Pro (10-core, 16 GB RAM) running Python 3.11 with NumPy 1.24.

F.3 Berry Inversion

The Berry (1994) inversion recovers the group-sector fixed effects $\delta_{g,s}$ that exactly rationalize the observed sectoral shares. Taking logs of the choice probability ratio between sector s and the outside option, and rearranging:

$$\delta_{g,s} = \sigma_\varepsilon \big[\ln \hat{\pi}_{g,s} - \ln(1 - \text{LFP}_g) \big] - \omega \ln w_s^{(0)} - \bar{z}_{g,s}$$

where $\hat{\pi}_{g,s}$ is the unconditional share of group g employed in sector s, $(1 - \text{LFP}_g)$ is the share choosing non-participation, $w_s^{(0)}$ is the baseline wage, and $\bar{z}_{g,s}$ is the mean productivity draw of group g in sector s (centered relative to the cross-group average). The fixed effects are normalized so that $\delta_{\text{MI,LE}} = 0$. This ensures that the baseline equilibrium exactly replicates both sectoral shares and participation rates.

F.4 Numerical Robustness

The solver's fixed-point iteration with dampening ($\lambda = 0.3$) ensures stable convergence across all tested parameter configurations. The sensitivity analysis in Appendix B systematically varies each calibrated parameter over a wide range (e.g., $\sigma_\varepsilon \in [0.5, 2.0]$, $\kappa^- \in [0.2, 0.6]$, $\kappa^+ \in [0.15, 0.45]$), confirming that the main findings are robust to input uncertainty and not driven by knife-edge calibration.