Motherhood Penalty in Labor Output: Evidence from Physicians*

Danyelle Branco Bladimir Carrillo Luiz Felipe Fontes Laisa Rachter

January 28, 2026

Abstract

This paper provides novel evidence that a motherhood penalty in earnings can arise through variation in labor output within the same career, even in the absence of reductions in labor supply or shifts into lower-paying jobs. We study this among female physicians in Brazil, where we can track motherhood, compensation, and detailed measures of output over time. Using a stacked event-study design, we find that mothers' earnings from medical services decline permanently following childbirth, driven mainly by reduced case volume and, to a lesser extent, a shift toward slightly lower-priced cases. The motherhood penalty in output is not explained by changes in working hours, broader career moves, or variation in how well or how intensively physicians treat their patients. Instead, it is driven by a series of within-job adjustments, such as mothers compressing their spatial and temporal availability, stepping back from complex, coordination-intensive procedures, and shifting toward lower value-added tasks within hospitals. These changes reshape physicians' opportunities to generate billable output, leading to persistent reductions in clinical production and, consequently, earnings.

JEL codes: *J13, J16, J24, J22, I10*

Keywords: *motherhood, child penalty, healthcare, physicians, labor productivity*

*Contact information: Branco: Sao Paulo School of Economics-FGV, Sao Paulo, SP, Brazil, 01332-000 (e-mail: danyelle.branco@fgv.br). Carrillo: Sao Paulo School of Economics-FGV, Sao Paulo, SP, Brazil, 01332-000 (e-mail: bladimir.carrillo@fgv.br). Fontes: Insper - Institute of Education and Research, Sao Paulo, SP, Brazil, 04546-042 (e-mail: luiz.fontes@insper.edu.br). Rachter: Inter-American Development Bank, Brasilia, Brazil, 70800-400 (e-mail: laisarachter@gmail.com). We also gratefully acknowledge financial support from the Brazilian National Council for Scientific and Technological Development (CNPq —# 04/2021). We thank Ana Balsa, Bruno Ferman, Fernanda Estevan, and Lucas Finamor, as well as participants at various seminars, for helpful conversations and comments on this project. We are solely responsible for this paper's contents. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

Childbearing is one of the most consequential events in women's careers,[1] with numerous studies documenting a motherhood penalty in the form of permanently reduced earnings.[2] Existing work on the causes of this penalty typically focuses on employment exits, reduced hours, and broad career moves into lower-paying jobs (Bertrand et al., 2010; Goldin, 2014; Kleven et al., 2019a; Cortés and Pan, 2023). In this paper, we argue that an important and largely overlooked component of the motherhood penalty can stem from changes in the volume and composition of tasks performed at work. In many professions, especially in highly skilled ones such as law, medicine, and financial advising, compensation and promotions are closely tied to what workers produce. If, following childbirth, women reduce their output or shift toward lower-remunerated forms of output, earnings losses can emerge and persist even when women keep the same job and hours worked.

Establishing the role of changes in how much and what women produce in shaping motherhood penalties is empirically challenging because datasets that jointly track motherhood, compensation, and detailed measures of output for many workers over time are rarely available. In this paper, we address this challenge in the context of physicians, a setting in which we can directly measure the quantity, quality, and type of output produced. Medicine is a highly specialized profession in which large career changes are relatively rare and costly, making within-job adjustments particularly salient. Indeed, using data from Brazil, we document that motherhood does not affect female physicians' labor supply, nor does it increase their likelihood of making broader career changes, in sharp contrast to evidence for broader populations. Yet we find significant and permanent losses in compensation that are largely explained by reduced case volume and, to a lesser extent, by a shift toward slightly lower-priced cases. The key message of this paper is that a motherhood penalty in earnings can arise even without reductions in labor supply or shifts into lower-paying jobs, through variation in output within the same career.

Our baseline sample comprises physicians working in hospitals within Brazil's Unified Health System (SUS), the largest publicly funded healthcare system in the world.[3] This setting is particularly relevant from a policy perspective, as motherhood penalties among physicians may have important consequences for patient welfare. We assemble novel administrative data that track physicians over time, combining information on payments received for medical services, detailed measures of the volume and composition of cases treated, their complete labor supply in healthcare facilities, and the exact timing of maternity leave spells. Our research design is an event study centered on the timing of birth, combined with a matching procedure that assigns placebo births to

[1] Motherhood is one of the main determinants of the gender income gap. As other determinants have become less important over time, the share of the gap explained by motherhood has increased from 40% in 1980 to 80% in 2013 (Kleven et al., 2019a).

[2] See, for instance, Angelov et al. (2016); Kleven et al. (2019a,b); Andresen and Nix (2022); Kleven et al. (2023); Cortés and Pan (2023); and Aguilar-Gomez et al. (2026).

[3] SUS is a single-payer, tax-funded system that provides universal health coverage to the population. In this context, physicians' remuneration is determined by a nationally standardized fee-for-service schedule. As such, differences in earnings primarily reflect variation in the volume and composition of services delivered, and are not the result of employer-specific pay policies. The next section provides further details about the institutional context.

observably similar nonmothers. To avoid biases from "forbidden" comparisons that could arise due to heterogeneity in treatment timing, we implement a stacked difference-in-differences framework (Gardner, 2022; Roth et al., 2023).

Our first set of results shows that motherhood entails substantial and lasting declines in physicians' revenues from hospital services. Total payments drop by roughly two thirds at the event date and then rebound only partially over the following semesters. In the long run, mothers' payments remain about 20 percent below those of comparable nonmothers. A natural explanation would be that mothers reduce their labor supply, but we find no evidence of such adjustments when using comprehensive data on physicians' work hours. In addition, when we scale payments by total work hours, the long-run penalty remains relatively similar. Thus, labor supply plays no role in explaining the payment decline.

We then open up the black box of compensation by decomposing payments into two key components: the volume of inpatient cases and payments per case. Both margins respond to childbirth. The number of cases falls steeply around the event and recovers only partially, converging to a long-run level about 14 percent below the counterfactual. Such a decline remains stable when we measure cases per hour devoted to hospital activity, reinforcing the absence of effects on labor supply. At the same time, payments per case decline modestly but persistently, by roughly 4 percent in the medium and long run. Using a simple decomposition, we show that the drop in case volume accounts for about 72 percent of the long-run decline in payments, while lower payments per case explain an additional 22 percent. The remaining 6 percent reflects a residual primarily capturing the extensive margin of payments, which is negligible in our data. Taken together, these results indicate that the motherhood penalty in physicians' earnings arises overwhelmingly from a sustained reduction in billable production, with a smaller contribution from a shift toward slightly lower-priced cases.

A key question is why output falls if hours do not. We examine several potential mechanisms. First, we investigate whether motherhood worsens the quality of care or changes treatment intensity. Drawing on risk-adjusted mortality and readmission rates, as well as measures of procedures, spending, and length of stay per case, we find no evidence of deteriorating quality or heightened intensity of care. Effects are close to zero and statistically insignificant across all horizons. These findings rule out explanations based on reductions in clinical performance, which could trigger hospital responses that limit mothers' exposure to cases, or on mothers spending substantially more time per case.

Second, we study how motherhood reshapes physicians' availability in space and time. Childcare and household logistics may increase the costs of travel, on-call duties, and irregular schedules, raising the value of predictability and local concentration. Consistent with this idea, the number of hospitals mothers practice declines by roughly 7.5 percent in the long run, and the number of distinct days worked per week falls by about 3.5 percent, with offsetting increases in hours per hospital and per day of activity. We also document a reallocation of work away from weekends toward regular weekdays. These patterns suggest that mothers compress their activity into a narrower set of

sites and time slots, which plausibly narrows the set of potential cases for which they are effectively available.

Third, we analyze changes in the composition of tasks within hospitals. A substantial share of hospital production involves complex, coordination-intensive procedures: major surgeries, high-technology interventions, and operations that require carefully scheduled rooms and teams. We find that motherhood reduces physicians' participation in these cases. The probability of performing any surgery falls, and the declines are larger for elective and high-technology procedures, on the order of 5–10 percent relative to baseline. Mothers also narrow the range of diagnoses they treat, suggesting a reduction in clinical breadth. At the same time, we observe shifts toward lower value-added responsibilities. The share of time allocated to non-clinical duties rises by almost one percentage point from a baseline of 9 percent, and in the years immediately following childbirth mothers are more likely to serve as collaborators rather than lead physicians on inpatient cases.

Finally, we assess the role of broader career changes. In principle, mothers could respond to the demands of childcare by changing occupations, switching employers, moving out of hospital-based care, or shifting their payer mix away from the public system. We find little evidence for these channels. Mothers become less, not more, likely to change medical occupations, and we do not detect increases in the probability that they move to a new facility. The share of work hours in hospitals remains roughly constant, and reliance on the public fee-for-service system does not decline.

To quantify the contribution of the different mechanisms, we follow a simple regression-based mediation logic, comparing baseline raw estimates and estimates adjusted for different sets of mechanisms. Controlling for the availability-dispersion variables attenuates the coefficient by roughly two-thirds, while controlling for changes in task composition attenuates the effect by almost one half. When we include both blocks jointly, the estimated long-run effect becomes small and statistically indistinguishable from zero. In contrast, adding controls for quality and intensity of care, hospital work hours, or for broader career changes leaves the estimates essentially unchanged. Taken at face value, these patterns suggest that most of the reduction in output can be accounted for by more concentrated availability across sites and days and by shifts away from complex, high-value-added clinical tasks.

Our paper speaks to the literature on motherhood and women's careers. A large body of work documents that women experience persistent earnings losses after childbirth, one of the most robust findings in labor economics and replicated across a range of settings (e.g. Angelov et al., 2016; Kleven et al., 2019a, 2023; Cortés and Pan, 2023; Aguilar-Gomez et al., 2026). By contrast, our understanding of the underlying mechanisms remains much more limited (Kleven et al., 2021; Andresen and Nix, 2022), and the leading explanation emphasizes reduced labor supply and broad career moves into lower-paying jobs.[4] We depart from this focus by providing novel evidence on the importance of changes in the volume, type, and value of output produced within the same occupation

[4]Kleven et al. (2021) and Andresen and Nix (2022) document motherhood penalties in earnings for non-birthing mothers in adoptive families, thereby ruling out biological factors related to childbirth as a relevant channel for these earnings losses.

and institutional environment. In this sense, our paper relates to recent work showing that mothers tend to have lower labor output than nonmothers or male workers within specific professional settings (Azmat and Ferrer, 2017; Gallen, 2024; Kim and Moser, 2025). We contribute to this literature by directly estimating the motherhood penalty in output and connecting it to changes in compensation. To our knowledge, we are the first to decompose the overall decline in compensation into reductions in billable production and changes in payments per unit of output, and show that the bulk of the earnings penalty can be accounted for by reduced output. We also move beyond establishing an output penalty to unpacking its drivers. We document a series of within-job adjustments, such as mothers compressing their availability to produce output, stepping back from complex, coordination-intensive procedures, and shifting toward lower value-added tasks. These mechanisms help reconcile our findings with previous evidence that mothers demand greater work flexibility (Bertrand et al., 2010; Goldin, 2014). However, prior literature typically rationalizes this demand as a cause of career moves into more flexible occupations, sectors, or firms that offer lower pay. Our results indicate that, in some settings, women may instead solve the flexibility problem through within-job adjustments, yet still face substantial earning losses when greater flexibility is achieved at the expense of exposure to valuable tasks within the firm.

Our paper also contributes to a longstanding literature in health economics that studies the determinants of the dispersion in the intensity and quality of care across physicians (e.g. Chandra and Staiger, 2007; Doyle Jr et al., 2010; Clemens and Gottlieb, 2014; Currie et al., 2016; Abaluck et al., 2016; Silver, 2021; Chan et al., 2022).[5] We add a new dimension to this literature by asking whether the transition into motherhood changes how well and intensively female physicians treat their patients. This question is particularly salient given descriptive evidence that motherhood can impair sleep, mood, and cognitive functioning, and that such constraints have been linked to physician clinical performance (e.g. Trockel et al., 2020). Using risk-adjusted mortality and readmission outcomes, as well as rich measures of treatment intensity, we find no evidence that motherhood changes either the quality or the intensity of care that patients receive. These findings are relevant in light of the "feminization of medicine" and the corresponding growing interest in the practice style and performance of female physicians (Tsugawa et al., 2017; Doyle Jr, 2020; Cabral and Dillender, 2024).[6]

The remainder of the paper is organized as follows. Section 2 provides background on the institutional setting and describes the data, sample, and variable definitions. Section 3 outlines our baseline research design. Section 4 examines the impact of motherhood on compensation and output volume. Section 5 explores mechanisms underlying the impact on output volume. Finally, Section 6 concludes.

[5] The literature has examined a wide range of factors underlying this heterogeneity, including place effects, peer effects, training, experience, skills, and financial incentives.

[6] Over the past century, women have transitioned from near exclusion from medical education to making up most new physicians across several countries, a trend commonly labeled the "feminization of medicine" (e.g. Phillips and Austin, 2009).

2 Setting and Data

This section describes the institutional setting and the data sources, samples, and variable definitions.

2.1 Institutional Setting

Brazilian Healthcare System. Following the enactment of the 1988 Constitution, Brazil established universal access to healthcare as a fundamental right, giving rise to the Unified Health System (SUS). This single-payer, tax-funded system ensures access to preventive and curative care via publicly managed institutions and a network of private facilities. Establishments in the SUS network provide healthcare services at no direct cost to patients. Instead, they are reimbursed by the federal government based on the volume and type of procedures performed, according to nationally standardized reimbursement rates established by the Ministry of Health. These rates also determine the compensation paid to healthcare professionals for clinical procedures. Individuals who seek care outside the SUS typically do so through private health insurance. In practice, most Brazilians rely on the SUS, while a minority (about 25%) hold private plans (Paim et al., 2011).

Physicians' labor market. Physicians in Brazil engage under various contractual arrangements, such as locum tenents, civil servants, regular employment contracted staff, or independent providers, often holding multiple jobs simultaneously. Within hospitals, the most common arrangements are standard employment contracts and independent contracting, which gives both physicians and hospitals considerable latitude in structuring physicians' work. It is common for physicians to maintain affiliations with several hospitals, even when not participating in their daily routines, allowing them to receive on-call requests and participate in specific cases.

Physicians may assume different roles within hospitals. A physician leading a case is responsible for diagnosing the patient's condition and selecting the appropriate treatment path. They deliver or oversee the main intervention, monitor recovery, and adjust treatment plans as needed, while supervising other members of the healthcare team. Physicians may also participate in supporting roles, reporting to the leading physician. In these cases, they may perform complementary activities from other specialties or carry out tasks within the same area under the leader's coordination; for example, managing an underlying clinical condition, conducting a diagnostic procedure, or assisting in surgical or therapeutic interventions. Within SUS hospitals, all such activities (primary or supporting) are compensated according to the national reimbursement rates, with the leading physician typically receiving the larger share. Beyond direct patient care, physicians often engage in non-clinical tasks, such as administrative work, coordinating supply requests, maintaining hospital records, and contributing to the organization of clinical services. These functions are essential to the daily operation of hospitals but generally fall outside formal payment structures.

2.2 Data

We use comprehensive data covering the universe of physicians in Brazil. These data draw on multiple administrative registries and include physicians' production and compensation within SUS hospitals, their complete labor supply in healthcare facilities, detailed professional and background characteristics, and all recorded instances of maternity leave. Our analysis focuses on female physicians with a positive caseload in SUS hospitals at baseline, whom we follow from 2008 through 2024.

2.2.A Data Sources and Variables

We gather data on the universe of all hospital admissions within SUS from the Hospital Information System (*Sistema de Informações Hospitalares*, SIH). SIH provides rich information on each admission, including patient characteristics (such as age and sex), patient mortality, admission and discharge dates, comorbidities and diagnoses coded according to the International Classification of Diseases (ICD-10), and detailed data on the procedures performed. For each admission, the dataset identifies one main procedure and, potentially, several secondary procedures. We observe the unique identifier of the healthcare professional who performed each procedure and the corresponding reimbursement for medical services paid following the national SUS fee schedule. We use SIH data to examine female physicians' compensation, case volume, practice characteristics (e.g., whether they perform surgeries), and performance indicators such as patient mortality and patients' average length of stay. A key concern is that variation in patient outcomes reflects not only differences in physicians' quality and practice style but also differences in the types of patients they treat. We implement a two-step risk-adjustment procedure to net out variation driven by observable patient characteristics.[7]

We have access to detailed background and human resource information for the universe of physicians in Brazil from the National Registry of Health Establishments (*Cadastro Nacional de Estabelecimentos de Saúde*, CNES). CNES is a rich monthly administrative database available since 2008 that covers every health facility in Brazil and its corresponding workforce. A major strength of these data is their universal coverage and high frequency: if a professional joins or leaves the staff of any health establishment, the facility is required to report this change to the system by the following month. Key professional variables include each physician's occupation, contracted hours, and the identifiers of all establishments where they provide inpatient or outpatient services. We use these

[7] In the first step, we regress a given outcome variable, y_{it}, for patient i in period t on the vector of patient characteristics, $\mathbf{X}_{it}$:

$$y_{it} = \beta \mathbf{X}_{it} + \mu_{jt} + \xi_{it},$$

where μ_{jt} represents a set of physician $\times$ period indicators. The vector of patient characteristics includes age, gender, race, zip code of residence, as well as a set of admission-level comorbidities. In the second step, we calculate the risk-adjusted outcome as follows:

$$y^*_{it} = y_{it} - \hat{\beta}\mathbf{X}_{it} \equiv \mu_{jt} + \xi_{it}.$$

Later, we collapse the residualized outcomes at the physician-by-period level and weight all regressions by cell size to account for differences in the precision with which the cell means are estimated.

data to study changes in physicians' labor supply, including total hours worked, the allocation of hours across different types of establishments, and the share of hours devoted to SUS-funded care.

Additional background information on physicians comes from the Federal Council of Medicine (CFM) and the National Medical Residency Commission (CNRM). The CFM is responsible for issuing and updating physicians' professional licenses, while the CNRM maintains records of residency titles granted by its affiliated programs. The primary purpose of these registers is to monitor and enforce compliance with regulatory standards for medical practice. These datasets collectively provide information on physicians' demographic characteristics, medical school attended, year of medical licensure, residency programs completed, and other credentials relevant for clinical practice. We use these data to identify a physician's gender and ensure that mothers and nonmothers are comparable in terms of key background characteristics.

Finally, we also use information from RAIS, Brazil's nationwide employer–employee administrative registry. These data track all formal employment relationships in the country, recording the beginning and end dates of each employment spell, the characteristics of employers and employees, and the exact dates of work absences due to workplace accidents, maternity leave, and other legally registered leave episodes. We use these data to identify the occurrence and timing of maternity leave episodes among female physicians.

2.2.B Baseline Sample and Sample Restrictions

Our initial sample consists of female physicians with some recorded inpatient activity in the hospital admissions data during the study period. We identify childbirth events in this sample based on the date at which a physician begins a maternity leave spell. For physicians with multiple spells, we retain only the earliest one. We observe 13,093 mothers in the data. Because our goal is to follow mothers for a long window before and after childbirth, we restrict the sample of mothers to those whose childbirth events occur between 2012 and 2018, and require that the physicians have positive cases within SUS hospitals in the year before the birth. This leaves us with 4,815 mothers in the analysis sample. Appendix Figure A.1 depicts a histogram of the timing of childbirth events in this sample, which closely mirrors the seasonal pattern of births in Brazil. We also restrict the set of nonmothers to be drawn from the same underlying population of female doctors providing care within SUS hospitals over the same period, yielding a pool of 56,692 physicians.

Appendix Table A.1 reports baseline descriptive statistics for mothers and nonmothers in the final sample. The table contains a broad set of covariates. Here we highlight just a few salient facts. One year before childbirth, mothers have on average 8 years of experience in medicine and 76% already hold an advanced specialization title.[8] They work, on average, 52 hours per week, are affiliated with two hospitals, and provide inpatient care on about 5.3 days of the week. The vast majority of their work time is devoted to SUS patients (88%) and to hospitals (74%). The same table shows that mothers and the full pool of nonmothers are overall quite similar along most observed

[8] Advanced specialization titles are formal residency titles registered with CNRM. There is also an informal path to medical specialization in Brazil. After six years of practical experience in an area of care, physicians can take an exam to be considered a specialist in that area. If they pass the exam, they receive the specialist title.

dimensions. The main difference is experience: naturally, mothers who give birth within our study window are younger than the broader set of nonmothers, whose average experience is 16 years. Having a specialization title at baseline is also noticeably more common among mothers, largely reflecting the fact that older physician cohorts in Brazil were much less likely to complete formal residency programs.

3 Research Design

Our empirical strategy is an event-study comparing mothers to similar nonmothers before and after birth events, following prior work on the effects of motherhood (e.g. Kleven et al., 2019a). To avoid biases from "forbidden" comparisons that arise in two-way fixed-effects models with staggered treatment adoption, we implement a stacked difference-in-differences framework (Gardner, 2022; Roth et al., 2023).

Control group. We begin by assigning placebo birth dates to female physicians who do not give birth during the study window, our control group. To do so, we match each mother to one nonmother physician and assign her the mother's birth quarter as a placebo birth date. Changes in outcomes for nonmothers around this placebo event trace the counterfactual path that mothers would have followed absent childbirth. To ensure a credible counterfactual, we select nonmothers with similar experience and professional background as the mother, and who were active in hospitals in the 12 months prior to the event. We implement coarsened exact matching to ensure that observations match exactly on year of medical licensure, licensing region, and indicators for holding an advanced specialization title, surgical practice, and graduation from a top-ranked medical school.[9] Appendix Table A.1 documents that mothers and placebo mothers are highly comparable across a very broad set of characteristics. The balance extends well beyond the covariates used in matching: the two groups also look nearly identical in measures of baseline inpatient production and compensation, labor supply, output mix, and quality and intensity of care.

Analysis sample. We construct separate panel datasets for each cohort of physicians experiencing a birth (or placebo birth) in a given quarter and then stack these subsamples into a single dataset to form the final analysis sample. We estimate event-study specifications in the stacked sample, including a full set of cohort fixed effects interacted with both unit and time fixed effects. This setup ensures that identification leverages comparisons between treated units and "clean" controls (i.e., nonmothers) within each cohort, addressing the pitfalls of two-way fixed-effects models in staggered designs. For estimation, we aggregate outcomes to semesters to obtain a parsimonious set of time-to-event effects with greater statistical precision. We also estimate the models on a fully balanced sample of physicians before and after the event. The longer the event-time window, the more observations are lost. To reduce this loss while capturing meaningful long-run dynamics, our

[9]This approach follows several recent applications of difference-in-differences with matching in labor settings (e.g. Goldschmidt and Schmieder, 2017; Jaravel et al., 2018; Smith et al., 2019; Arnold, 2022).

event window spans 5 semesters before and 12 semesters after childbirth. We later demonstrate that the results are robust to alternative choices of window length, including specifications with longer pre- and post-periods.

Estimating equations. We use the following event-study specification to estimate the effect of motherhood:

$$\begin{aligned} y_{itc} = \alpha + \sum_{\tau \neq -2} \beta_\tau \text{Motherhood}_{ic} \times \mathbb{1}[\text{semester since childbirth} = \tau] \\ + \ \mu_{ic} + \ \lambda_{tc} \ + \eta_{itc}, \end{aligned} \tag{1}$$

where y_{itc} is the outcome of interest for female doctor i in semester-year t and event cohort c. Motherhood is a dummy variable that equals one for mothers and zero for placebo mothers. The terms μ_{ic} and λ_{tc} represent doctor×cohort and semester-year×cohort fixed effects. The coefficients $\{\beta_\tau\}$ trace the path of outcomes around the event and are identified relative to the omitted bin $\tau = -2$, two semesters before childbirth. Following standard practice in the literature, we use a pre-pregnancy period as the omitted category to avoid potential biases from the effects of pregnancy per se. The residual term, η_{itc}, is clustered at the doctor level. We later show that inference is robust to alternative clustering choices.

To summarize magnitudes in table format and perform specification checks parsimoniously, we also estimate a reduced version of (1):

$$\begin{aligned} y_{itc} = \alpha + \beta_0 \text{Motherhood}_{ic} \times \text{Short-Run}_t + \beta_1 \text{Motherhood}_{ic} \times \text{Medium-Run}_t \\ + \beta_2 \text{Motherhood}_{ic} \times \text{Long-Run}_t + \ \mu_{ic} + \ \lambda_{tc} \ + \gamma \text{Pregnancy}_{itc} + \eta_{itc}, \end{aligned} \tag{2}$$

where the interaction between Motherhood and Short-Run captures the effects up to two semesters following the childbirth event. The variable Medium-Run is an indicator for the 3–8 event-time window. Long-Run is an indicator for semesters beyond that horizon and summarizes the persistent effects of childbirth. We control for $\text{Pregnancy}_{itc} \equiv \text{Motherhood}_{ic} \times \mathbb{1}[\text{one semester prior to childbirth}]$ to ensure that all treatment effects are evaluated relative to a prepregnancy period. The remaining variables and coefficients are identical to those in equation (1).

Identification. A doctor's decision to have a child may not be exogenous. If it is driven by relatively fixed traits—such as career orientation or preferences for work–family balance—this does not threaten our design, which exploits within-physician variation. For a causal interpretation , we rely on a standard parallel-trends assumption: mothers and matched placebo mothers would have followed similar trends in the absence of childbirth.[10] The main threats to identification stem from unobserved, time-varying determinants of childbirth that also influence labor market outcomes. Examples include changes in family circumstances (for instance, a partner's promotion that relaxes income constraints) or shifts in the doctor's role at the hospital that simultaneously make childbirth

[10]Since placebo mothers are assigned based on matching, identification in this setting relies on conditional parallel trends between mothers and nonmothers, a weaker assumption than requiring unconditional parallel trends (Roth et al., 2023).

more attractive and alter her workload. However, most such factors would likely affect outcomes prior to the birth event and show up as divergent pre-trends. We directly assess this threat using the event-study equation (1). As documented throughout the paper, pre-event coefficients are flat and tightly centered around zero in our data. Mothers and controls diverge only after childbirth, and in a very sharp way. For time-varying unobservables to generate this pattern, they would need to arise abruptly and systematically precisely at the event date, and only for actual mothers. We view this scenario as less likely.[11]

Identification concerns are further mitigated by the strong baseline similarity between treated and placebo mothers in our sample, which makes the design robust to time-varying shocks that depend on predetermined physician characteristics. We also show that our estimates remain very stable when we incorporate more flexible time controls, including time effects specific to each matching pair. Finally, identification in this setting also requires an assumption of no anticipatory effects. This assumption would be violated if female physicians adjust their labor supply or practice patterns in advance because they expect to have a child. Pre-trends provide a direct test of this, and our flat estimates before the event strongly support the assumption.

4 Effects of Motherhood on Compensation and Output

This section investigates how the transition into motherhood affects physicians' revenues from medical services, the volume of cases they attend, and the payments received per unit of output.

4.1 Payments for Medical Services

Figure 1 presents event-study estimates for the impact of motherhood on physicians' payments for medical services. Panel A focuses on the extensive margin, reporting effects on an indicator variable that equals one if the physician receives any payment in a given semester. Panel B turns to the log of payments, calibrated to account for the extensive margin following Chen and Roth (2024).[12] In both panels, coefficients in the prepregnancy period are essentially zero with tight confidence intervals, supporting the validity of the design. The first deviations from zero emerge in the semester immediately preceding childbirth, followed by a sharp discontinuity in the childbirth semester.

In Panel A, the probability of receiving any payment falls by about 15 percentage points at the event date. This decline is short-lived: by the fourth semester after childbirth, the effect is statistically indistinguishable from zero, with point estimates subsequently converging toward zero. This muted, temporary response indicates that motherhood does not prompt physicians to exit medical

[11] Even though the parallel-trends assumption is ultimately untestable, Kleven et al. (2019a) show that child penalty estimates from difference-in-difference models remain essentially stable when one tries to correct for potential endogeneity using popular IV approaches.

[12] Because the extensive-margin effect is small and transitory, we minimize its influence on the log transformation. Specifically, we set log(0) equal to the log of the 1st percentile positive value. We later show that our results are highly robust to alternative calibrations.

practice within SUS hospitals. Once we incorporate intensive-margin variation, Panel B reveals a markedly more persistent effect. Total payments collapse by roughly two thirds at childbirth and bounce back quickly over the subsequent two semesters, but only to a level that remains about 25 percent below that of placebo mothers. Beyond that point, there is little additional recovery: from the fourth year after childbirth onward, the penalty stabilizes around 19–20 percent.

Panel A of Appendix Figure A.2 shows that our estimate of the long-run penalty is very similar when we rely only on intensive-margin variation or adopt alternative strategies for handling zero payments. This suggests that our results are not an artifact of the calibration and confirms that the bulk of the compensation decline is not driven by the extensive margin. In addition, we later show that the estimated long-run penalty remains stable over extended horizons. Taken together, these findings indicate that the transition into motherhood entails substantial and persistent economic losses for female physicians, consistent with evidence from other contexts. To put the magnitude of our estimate in perspective, it is comparable to the long-run child penalty in earnings documented for Scandinavian countries (Kleven et al., 2019b).

Labor supply. A leading explanation for the persistent decline in physicians' compensation is a reduction in labor supply. The preceding results indicate that mothers do not exit medical practice altogether, but they could still scale back their workload. To assess this possibility, we use data from physicians' full labor supply in healthcare and estimate average effects on the log of work hours using equation (2). Appendix Table A.2 presents the results. Across all time horizons, the estimates are small (all below 0.8%) and not statistically significant. The same table estimates effects on the log of payments per work hour. Unsurprisingly, this does not overturn the result on total payments: the decline remains statistically significant and large in the long run, at about 16 percent. Hence, the drop in physicians' revenue after childbirth is not driven by lower labor supply. This stands in sharp contrast to evidence for broader populations, where reductions in labor supply have been shown to account for a large share of the motherhood penalty in earnings.

4.2 Output Volume and Payments per Output

Compensation in this setting is tightly linked to how much physicians produce and which services they provide. Following childbirth, physicians may attend to fewer cases and shift toward services that are reimbursed at lower rates. Such production shifts can generate large and persistent payment losses even when total work hours remain stable. We separate these channels by estimating effects on both the volume of inpatient cases and the payments received per case. Figure 2 presents the results. Panel A reports event-study estimates for the log of cases. As before, coefficients in the prepregnancy period are flat around zero. The number of cases falls sharply in the event semester and then rebounds quickly over the next periods. The recovery is only partial, though. From about the second year onward, the gap settles at a new, lower level and remains roughly constant through the end of the event window. In the long run, we see a decline of about 14 percent in cases. Consistent with the evidence on payments, our results are virtually unchanged across alternative

calibrations of the extensive margin (Panel B of Appendix Figure A.2).

These findings suggest that a sizable share of the drop in payments reflects lower production along the intensive margin. One could still argue that, even if mothers' total labor supply is stable, they may reallocate time away from hospital-based care toward other types of medical work, such as private clinics or primary health care facilities. To investigate this, we exploit granular information on work hours and estimate effects on the log of cases per hour devoted to hospital work. Appendix Figure A.3 plots the event-study coefficients for log cases and log cases per hospital hour for physicians with positive cases. The two series are nearly indistinguishable, so the reduction in volume is not driven by reduced labor supply within hospitals, but by fewer cases per hour on duty. This pattern points to a decline in operational productivity (in terms of billable output), a margin largely unexplored in prior work.

Panel B examines the log of payments per case. Here the effects are smaller in magnitude but highly persistent. Coefficients in the prepregnancy period are again statistically indistinguishable from zero, while a modest decline appears already in the pregnancy semester and continues after childbirth. Over the medium and long run, payments per case are consistently about 4 percent lower for mothers than for nonmothers. This pattern indicates that, beyond treating fewer cases, mothers shift toward cases that are reimbursed at slightly lower rates.

4.3 Decomposing the Payments Decline

Having documented that both the volume of cases and payments per case decline after childbirth, we now quantify how much each margin contributes to the overall loss in physicians' revenue from hospital services. For physicians with positive cases (and thus positive payments), log payments satisfy the identity $\log(\text{payments}) = \log(\text{volume of cases}) + \log(\text{payments/volume of cases})$. Let Δz denote the long-run difference-in-differences effect of motherhood on outcome z. Since Δ is a linear operator, the long-run effect on log payments can be written exactly as the sum of the long-run effects on the log of cases and the log of payments per case. Our baseline estimate for the long-run penalty in payments is, however, based on a calibrated version of $\log(\text{payments})$ that also incorporates variation from the extensive margin—call it $\widetilde{\log}(\text{payments})$. We can decompose the impact on this outcome into the two components above and the residual term $R \equiv \Delta\widetilde{\log}(\text{payments}) - \Delta\log(\text{payments})$:

$$\Delta\widetilde{\log}(\text{payments}) = \Delta\log(\text{volume of cases}) + \Delta\log(\text{payments/volume of cases}) + R, \qquad (3)$$

where one can show that the relevance of R largely depends on the baseline share of female physicians with no payments for medical services and the long-run treatment effects on the likelihood of receiving positive payments, both of which are very close to zero in our setting.[13]

We use the formulation in (3) to decompose the estimated decline in payments into the share

[13] Let π_0^M and π_0^P denote the baseline shares of mothers and placebo mothers with positive payments, and let $\Delta\pi^M$ and $\Delta\pi^P$ denote the within-group changes in these shares between the baseline and long-run periods. Define the treatment effect on the extensive margin as $\Delta\pi \equiv \Delta\pi^M - \Delta\pi^P$. Let ℓ_1^M and ℓ_1^P be the long-run averages of $\log(\text{payments})$ among physicians with positive payments in each group, and let $\Delta\ell^M$ and $\Delta\ell^P$ be the corresponding within-group changes over

driven by reduced volume of cases, the share driven by lower payments per case (both computed exploiting only intensive-margin variation), and the residual component R. Figure 3 presents the results of this decomposition. The volume of cases accounts for the bulk of the long-run decline in payments: 72.4 percent is explained by lower inpatient production. A further 21.8 percent is explained by a downgrading of output (i.e., lower payments per case). The residual component amounts to only 5.8 percent, consistent with earlier results indicating that the extensive margin plays a minor role. These shares line up closely with the magnitudes of the underlying long-run effects documented in the subsection above.

Taken together, these results show that the motherhood penalty in physicians' revenue arises overwhelmingly from a sustained reduction in billable inpatient activity, with a smaller but non-negligible contribution from a shift toward slightly lower-priced cases.

4.4 Robustness Checks

We now conduct a battery of robustness and specification checks to assess potential threats to our main findings.

Pre-trends. In the previous subsections, we showed that pre-treatment estimates consistently hover around zero. Appendix Table A.3 provides additional validation tests of the parallel pre-trends. Specifically, it reports two sets of F-tests for joint significance. The first set uses our baseline event window, in which the pre-treatment period runs up to 5 semesters before the event. The second set extends the window to include three additional pre-event semesters. In both cases, and for each of the three outcomes, we never reject the null hypothesis that all pre-treatment coefficients are jointly equal to zero. The similarity of the F-statistics across the two windows indicates that mothers and nonmothers followed parallel trends over a long pre-event horizon.

Stability to time-specific trends. Our baseline specification includes cohort-specific time fixed effects. As a test of robustness to differential time trends, we estimate an alternative specification that replaces these with time-by-matching-strata fixed effects. In this setup, we allow for fully flexible time paths for every combination of baseline physician characteristics used in the matching, so that our estimates are robust to any time trends tied to those covariates. Appendix Table A.4 reports the results. Panel A shows that the estimated long-run penalty in payments is highly stable across specifications: it changes only from -0.194 in the baseline to -0.192 when we include stratum-specific time effects. At the same time, the R^2 increases substantially, from 0.58 to 0.70, indicating that the additional time controls absorb a large share of the variation in the outcome without altering the estimated treatment effect. Following the logic in Oster (2019), this pattern suggests that the

time. Under the calibration $\log(0) = \log m$, simple algebra implies that

$$R = \Delta\pi\,(\ell_1^M - \log m) + (\pi_0^M - 1)\,\Delta\ell^M + (1 - \pi_0^P)\,\Delta\ell^P + \Delta\pi^P(\ell_1^M - \ell_1^P).$$

In our data, π_0^M and π_0^P are equal to one and the changes $\Delta\pi$ and $\Delta\pi^P$ are close to zero, so R should be empirically small.

presence of unobservable time-varying factors would not materially change our conclusions. In particular, we can use Oster's framework to compute bounds for the estimated effect in the presence of omitted variables. Under standard parameter choices,[14] the implied identified set for the long-run effect lies in the interval $[-0.194, -0.189]$. Moreover, this interval remains comfortably away from zero for any reasonable value of the selection parameter governing the relationship between observables and unobservables. Panels B and C show a similarly high degree of stability for the long-run effects on log cases and log payments per case, respectively.

Longer-run effects. Our main results are estimated in a balanced panel that follows physicians up to 12 semesters after childbirth. One could worry that, over a longer horizon, the estimated treatment effects might eventually converge back to zero. We therefore examine the longer-run dynamics of the treatment effects when we extend the post-event window. Appendix Figure A.4 shows that this is not the case. The figure reports time-to-event treatment effects up to 18 semesters after childbirth and, for ease of interpretation, also includes a horizontal dashed line indicating the average long-run effect estimated via equation (2) in our baseline sample. The coefficients for the additional periods remain virtually unchanged relative to this benchmark, suggesting that the penalties are indeed permanent.

Excluding prior childbirths. In our empirical strategy, the treatment event for mothers is defined as the first childbirth observed within the study period. Similarly, the control group consists of women who do not experience any childbirth during this window. A potential concern is that some physicians may have had children before the start of our observation period, so that we may not always be capturing the very first birth (or the true absence of children). To assess the relevance of this issue, we use maternity-leave records going back to 2003 and exclude from the sample all physicians with any recorded maternity leave prior to the study period. Appendix Figure A.5 compares the long-run effects on our main outcomes in the baseline sample with those obtained under this restriction. The two sets of estimates are essentially identical, indicating that this concern is not empirically relevant. This stability is unsurprising, as the restriction removes only about 8 percent of the sample, roughly evenly split between treated and control physicians.

Inference. Finally, we assess the robustness of our inference to alternative clustering choices. Our main results cluster standard errors at the physician level to account for heteroskedasticity and serial correlation within doctors over time. As a robustness check, we also report two-way clustered standard errors at the physician and semester-year levels, allowing for arbitrary correlation both within doctors and across doctors exposed to common shocks in the same period. In addition, we test for clustering at the matching-pair level, which accommodates the dependence induced by the first-stage matching procedure (Abadie and Spiess, 2022), and at the municipality level to allow for spatially correlated shocks across physicians working in the same local market. Appendix

[14] As suggested by Oster (2019), we assume that incorporating omitted variables would raise the R2 by an additional 30 percent, and that unobserved and observed determinants are equally correlated with motherhood.

Figure A.6 plots the absolute t-statistics for the long-run effects on our main outcomes under each clustering scheme. Across the board, the statistics remain well above the 1.96 benchmark for statistical significance at the 95% level, and in practice they change very little relative to our baseline physician-level clustering. This evidence indicates that our standard errors are highly stable across different inference assumptions.

5 Mechanisms

The results so far indicate that motherhood leads to sizable and persistent declines in physicians' revenue from hospital services, driven primarily by reduced case volume and, to a lesser extent, by slightly lower payments per case. We have already ruled out reductions in labor supply as an explanation for this output decline. In this section, we investigate alternative channels, with a particular focus on how motherhood reshapes the way physicians practice medicine and organize their work. Given the large number of outcomes, we summarize the evidence in tables estimated using the reduced specification (2). Appendix Table A.5 reports that, for all mechanism outcomes and across alternative pre-treatment windows, we fail to reject the null that the pre-event coefficients are jointly equal to zero, supporting the validity of our design for this broader set of variables.

5.1 Quality and Intensity of Care

Motherhood has been linked to impairments in sleep, mood, and cognitive functioning (Orchard et al., 2023; Henry and Rendell, 2007; Swain et al., 1997). In medical settings, such constraints have been associated with higher rates of clinical errors and adverse patient outcomes (e.g., Trockel et al., 2020). If motherhood worsened the quality of care delivered by female physicians, hospitals might have responded by assigning them fewer and lower-stakes cases (Sarsons, 2022), a pattern that could help explain our findings.[15] To assess this hypothesis, we use the inpatient mortality rate among physicians' patients as our primary measure of care quality. To address concerns that mortality differences across physicians may reflect variation in patient case mix rather than quality of care, we use risk-adjusted mortality following the approach in Section 2.2. Column 1 of Table 1 presents our results. In the short, medium, and long run, the coefficients are essentially equal to zero, indicating no deterioration in care quality. Column 2 complements this evidence by investigating 30-day risk-adjusted readmission rates, which may capture complications and downstream health problems not fully reflected in mortality. Consistently, we find no relevant evidence of deteriorating performance. A small, marginally significant effect appears in the short run (about 3 percent of the baseline mean) but effects at other horizons are consistently close to zero and statistically insignificant.

A different mechanism within the domain of patient care that could still help explain our results is increased intensity of care. Following childbirth, female physicians might become more cautious or more involved with each case. For example, ordering more secondary procedures, monitoring

[15] Sarsons (2022) show that when patients die under the care of female surgeons, these surgeons are disproportionately penalized in subsequent case assignments, receiving fewer and less complex surgeries.

patients more closely, or being more reluctant to discharge. Such changes would raise the amount of time physicians invest in each case and, in turn, reduce overall output. In columns 3, 4, and 5 of Table 1, we examine changes in the intensity of care. We focus on the risk-adjusted number of procedures per case, spending per case, and length of stay. The estimates are small and statistically indistinguishable from zero across the board, indicating motherhood does not affect how intensively physicians treat the patients they see.

5.2 Availability Dispersion

We now investigate how motherhood reshapes physicians' availability in both space and time. Childcare and household logistics may raise the costs of travel, handoffs, and irregular schedules, increasing the value of predictable workdays and the cost of cross-site coordination. In response, mothers may choose to concentrate their activity within a narrower set of sites and time slots. This reduced dispersion of availability could narrow the set of cases to which mothers are effectively exposed. For instance, by limiting the network of colleagues with overlapping schedules, reducing referral opportunities, or decreasing participation in on-call work.

Table 2 summarizes how motherhood affects the availability of physicians across hospitals and days of work. Column 1 begins by examining the number of hospitals at which a physician practices. At baseline, female physicians are affiliated with about two hospitals. In the short run, motherhood leads to a statistically significant but small decline of 0.06 hospitals. This effect intensifies over time: in the long run, the number of hospitals falls by 0.15, or roughly 7.5 percent relative to baseline. Column 2 shows that work hours per hospital rise in parallel. The long-run effect is an increase of 1.4 hours per hospital, about 7.2 percent relative to a baseline mean of 20 hours. These patterns are consistent with total hours worked remaining roughly constant but increasingly concentrated in a smaller set of sites.

Column 3 investigates the number of distinct days per week on which a physician practices within hospitals.[16] At baseline, physicians work on about 5.1 days per week. Motherhood induces a sharp short-run decline in days worked, followed by a partial recovery and then a stable gap over time. The long-run estimate imply a persistent reduction of about 0.17 days per week. Column 4 shows a small offsetting increase in hours worked per day of activity, of similar magnitude in percentage terms, again pointing to a shift from dispersed to more concentrated work schedules. Table A.6 further explores how work is reallocated across specific days of the week. We find a significant long-run decline in the share of weekend days in physicians' weekly activity, driven especially by a reduction in Sunday work. This pattern is consistent with childcare gaps and more constrained household logistics on weekends. In parallel, we observe some reallocation toward regular weekdays. These patterns do not support a "rescheduling to low-demand days" explanation for our findings: weekends have fewer cases on average, so shifting away from them would, if anything, attenuate rather than amplify the observed volume losses.

[16] We construct this measure using the calendar dates on which the physician treats hospital patients in a given week. We then count the number of distinct days with at least one case. The baseline mean of this measure is equivalent to a standard full-time weekly schedule.

5.3 Composition of Tasks

If coordination and unpredictability become more costly after childbirth, physicians may reoptimize not only their availability, but also what they do. A significant share of hospital work consists of complex cases that require long, irregular hours or extensive on-call availability, such as major surgeries and high-acuity, high-technology procedures. If female physicians shy away from such cases following childbirth, their output should fall and they may be pushed toward lower value-added activities.

Table 3 examines how the composition of physicians' tasks within hospitals changes after childbirth. It first investigates, among physicians with positive clinical output, whether motherhood reduces their participation in complex cases. Column 1 reports that, in the long run, the probability of participating in any surgery falls by about 0.9 percentage points relative to a baseline mean of 38 percent. The decline becomes larger as we move to more demanding procedures. Column 2 considers an indicator equal to one if the physician takes part in elective procedures—typically more complex interventions that require scheduling specific rooms, teams, and time slots. Here, the long-run effect reaches a reduction of about 2 percentage points, or roughly 5 percent of the baseline level. Column 3 focuses on high-complexity procedures,[17] for which the long-run decline is on the order of 10 percent in relative terms. These results indicate that mothers gradually shift away from the most complex cases within hospitals. Column 4 turns to the clinical breadth of physicians' practice. If mothers restrict the range of conditions in which they engage following childbirth, this could further limit the set of cases for which they are effectively available and contribute to the overall decline in output. To investigate this, we estimate effects on the number of distinct diagnoses they treat. The long-run coefficient implies a meaningful decline of about 0.7 diagnoses or 5% of the baseline.

We next examine which activities mothers appear to substitute toward. Qualitative work suggests that female physicians often end up taking on a range of "invisible" or undervalued responsibilities within hospitals that consume time but receive little formal acknowledgment (Lim et al., 2021). Examples include ensuring compliance with institutional protocols, handling routine documentation, and organizing logistics and scheduling. Beyond administrative duties, women may be frequently pulled into supporting roles on inpatient cases led by other physicians. Motherhood could amplify the likelihood that female physicians shift into these roles. Measuring effects on such low-value tasks is empirically challenging, though, as they are rarely captured in standard administrative data. Our goal is not to identify every specific task to which physicians reallocate time, but to look for evidence on whether their hospital work seems to lean toward these lower value-added activities following childbirth.

Column 5 of Table 3 examines the share of cases in which the physician serves as a collaborator rather than as the lead doctor. In the short and medium run, we find a statistically significant increase in this supporting role: the share of cases in which the physician acts as a collaborator

[17] SUS classifies the portfolio of procedures offered within the Brazilian healthcare system into three categories according to their complexity: low, medium, and high. High-complexity procedures are those that require advanced technology, such as organ transplants, cardiac surgery, and treatment for patients with third-degree burns.

rises by about 1.5–1.7 percentage points, relative to a baseline mean of 53 percent. By the long run, however, this effect shrinks to around 0.4 percentage points and is no longer statistically significant, suggesting that the shift away from leadership roles is more pronounced in the years immediately following childbirth. Column 6 then considers non-clinical duties. In our context, health authorities record the total hours physicians devote to non-clinical activities, allowing us to measure the share of time allocated to such tasks. We find a persistent increase that amplifies over time: in the short run, involvement in non-clinical duties rises by about 0.4 percentage points relative to a baseline of 9 percent; in the long run, the effect nearly doubles to 0.8 percentage points. These results suggest that part of the decline in inpatient output following childbirth may reflect a reallocation of mothers' time toward low-value-added responsibilities.

5.4 Broader Career Changes

A prominent view in the literature is that earnings losses after childbirth reflect not only reduced hours and labor-force participation but also broad career adjustments, such as switching employers, sectors, and occupations (Goldin, 2014; Kleven et al., 2019a; Cortés and Pan, 2023). We have already shown that neither exit nor fewer hours explains the revenue losses we document; instead, they are driven primarily by lower billable output. In a highly specialized, licensed profession such as medicine, large career changes outside the field are relatively rare and costly. Still, one might rationalize the decline in inpatient volume as the by-product of broader career shifts within medicine. Mothers could, for example, pivot toward new medical occupations or employers that favor flexibility over output, reallocate effort away from hospitals toward other practice settings, or tilt their payer mix toward private patients and away from SUS fee-for-service.

Table 4 assesses these alternatives. First, mothers become less–not more–likely to change occupations: the probability falls by about 4 percentage points, and the decline persists in the long run. Second, we find no increase in the likelihood that a physician changes to a new health facility. Effects are negative in the short run and statistically indistinguishable from zero at longer horizons. Third, the share of work hours in hospitals remains stable following childbirth, so female physicians continue to allocate most of their time to hospital work ($\approx$ 75 percent). Finally, reliance on SUS does not fall; if anything, it edges up modestly. The share of hours female physicians devote to SUS increases by about 2 percentage points in the long run from a baseline of 90 percent.

In summary, motherhood does not lead physicians to switch occupations or jobs, or to reduce their participation in hospitals and within the SUS network. This indicates that, in our setting, the key adjustments operate through within-occupation and within-employer changes rather than through broad career transitions.

5.5 Relevance and Summary of Mechanisms

We now assess to what extent the mechanisms investigated above can account for the decline in output volume. We follow a simple regression-based mediation logic, comparing baseline and adjusted estimates in the spirit of Gelbach (2016). Specifically, we reestimate the effects on log cases

using equation (2) and then sequentially augment the specification with different blocks of mechanism variables. The extent to which the estimated effect is attenuated when a given block is added provides suggestive evidence on how much that mechanism can account for the observed output decline. All our specifications include hospital hours as a control to ensure that the mechanism covariates do not capture variation in labor supply.

Figure 4 summarizes these results. The first plot reports the baseline long-run coefficient on log cases, equal to -0.14. The second plot shows that simply adding hospital hours left the estimated effect unchanged, which is unsurprising given our earlier results. We then turn to the mechanisms for which we found substantial effects. Including the availability-dispersion measures shrinks the long-run coefficient on log cases from -0.14 to -0.051, suggesting that changes in where and when physicians are available account for a sizable share of the volume decline. Adding the task-composition block instead lowers the effect to -0.079, indicating a somewhat smaller but still meaningful role for the within-hospital composition of output. When we include both sets of mechanism variables jointly, the estimated long-run effect falls to -0.038 and is no longer statistically significant at the 95% level. Taken at face value, these estimates suggest that over two-thirds of the negative impact of motherhood on case volume can be explained by the combination of reduced dispersion of availability and shifts in task composition. As a sanity check, Appendix Figure 4 also reports specifications that add controls for broader career changes and for quality and intensity of care. Neither set of controls attenuates our main estimates, consistent with the absence of motherhood effects on these dimensions.[18]

Taken together, the evidence on mechanisms paints a coherent picture. Motherhood does not prompt female physicians to reduce labor supply or make broad career changes, nor does it affect how well or how intensively physicians treat their patients. Instead, it induces a series of adjustments to the mothers' schedules and output composition. Female physicians narrow their clinical scope, step back from complex, coordination-intensive procedures, shift toward lower-value-added tasks in hospitals, and compress the spatial and temporal dispersion of their availability. These changes reshape physicians' opportunities to generate billable output, leading to persistent reductions in production and, consequently, lower earnings. These patterns are consistent with prior work emphasizing mothers' demand for temporal flexibility, but they shift the relevant mechanism from between-job sorting to within-job assignment.

6 Concluding Remarks

This paper documents that motherhood can generate substantial and persistent earnings losses for women even when labor supply and broad career trajectories remain unchanged. Using comprehensive administrative data on Brazilian physicians and a stacked event-study design, we provide evidence of a significant motherhood penalty in earnings that arises primarily from reduced case

[18]This does not imply that these variables are unrelated to output. For instance, among placebo mothers, a 1–percentage-point increase in the risk-adjusted mortality rate is associated with roughly a 27% reduction in case volume (conditioning on physician and time-by-cohort fixed effects).

volume and, to a lesser extent, slightly lower payments per case. The primary drivers of these output declines are reduced spatial and temporal availability and a shift away from complex, coordination-intensive tasks toward secondary roles. We find no evidence of reductions in work hours, broader career changes, or changes in patient outcomes. These findings highlight the importance of within-job production shifts as a source of gender gaps in high-skill professions. More broadly, they underscore that even when working mothers solve the temporal-flexibility problem within their careers, they may still face significant economic losses if doing so reduces their opportunities to produce high-value output. Future work could assess whether similar mechanisms operate in other high-skill occupations and evaluate which institutional arrangements best reconcile temporal flexibility with the sustained production of high-value output.

Bibliography

Abadie, Alberto and Jann Spiess, "Robust post-matching inference," *Journal of the American Statistical Association*, 2022, *117* (538), 983–995.

Abaluck, Jason, Leila Agha, Chris Kabrhel, Ali Raja, and Arjun Venkatesh, "The determinants of productivity in medical testing: Intensity and allocation of care," *American Economic Review*, 2016, *106* (12), 3730–3764.

Aguilar-Gomez, Sandra, Eva O Arceo-Gomez, and Elia De la Cruz Toledo, "Inside the black box of child penalties: Unpaid work and household structure," *Journal of Development Economics*, 2026, *178*, 103554.

Andresen, Martin Eckhoff and Emily Nix, "What causes the child penalty? Evidence from adopting and same-sex couples," *Journal of labor economics*, 2022, *40* (4), 971–1004.

Angelov, Nikolay, Per Johansson, and Erica Lindahl, "Parenthood and the Gender Gap in Pay," *Journal of Labor Economics*, 2016, *34* (3), 545–579.

Arnold, David, "The impact of privatization of state-owned enterprises on workers," *American economic journal: applied economics*, 2022, *14* (4), 343–380.

Azmat, Ghazala and Rosa Ferrer, "Gender gaps in performance: Evidence from young lawyers," *Journal of Political Economy*, 2017, *125* (5), 1306–1355.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz, "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics*, July 2010, *2* (3), 228–55.

Cabral, Marika and Marcus Dillender, "Gender differences in medical evaluations: evidence from randomly assigned doctors," *American Economic Review*, 2024, *114* (2), 462–499.

Chan, David C, Matthew Gentzkow, and Chuan Yu, "Selection with variation in diagnostic skill: Evidence from radiologists," *The Quarterly Journal of Economics*, 2022, *137* (2), 729–783.

Chandra, Amitabh and Douglas O Staiger, "Productivity spillovers in health care: evidence from the treatment of heart attacks," *Journal of political Economy*, 2007, *115* (1), 103–140.

Chen, Jiafeng and Jonathan Roth, "Logs with zeros? Some problems and solutions," *The Quarterly Journal of Economics*, 2024, *139* (2), 891–936.

Clemens, Jeffrey and Joshua D Gottlieb, "Do physicians' financial incentives affect medical treatment and patient health?," *American Economic Review*, 2014, *104* (4), 1320–1349.

Cortés, Patricia and Jessica Pan, "Children and the remaining gender gaps in the labor market," *Journal of Economic Literature*, 2023, *61* (4), 1359–1409.

Currie, Janet, W Bentley MacLeod, and Jessica Van Parys, "Provider practice style and patient health outcomes: The case of heart attacks," *Journal of Health Economics*, 2016, *47*, 64–80.

Gallen, Yana, "Motherhood and the gender productivity gap," *Journal of the European Economic Association*, 2024, *22* (3), 1055–1096.

Gardner, John, "Two-stage differences in differences," *arXiv preprint arXiv:2207.05943*, 2022.

Gelbach, Jonah B, "When do covariates matter? And which ones, and how much?," *Journal of Labor Economics*, 2016, *34* (2), 509–543.

Goldin, Claudia, "A grand gender convergence: Its last chapter," *American economic review*, 2014, *104* (4), 1091–1119.

Goldschmidt, Deborah and Johannes F Schmieder, "The rise of domestic outsourcing and the evolution of the German wage structure," *The Quarterly Journal of Economics*, 2017, *132* (3), 1165–1217.

Henry, Julie D and Peter G Rendell, "A review of the impact of pregnancy on memory function," *Journal of clinical and experimental neuropsychology*, 2007, *29* (8), 793–803.

Jaravel, Xavier, Neviana Petkova, and Alex Bell, "Team-specific capital and innovation," *American*

Economic Review, 2018, *108* (4-5), 1034–1073.

Jr, Joseph J Doyle, "Physician characteristics and patient survival: Evidence from physician availability," Technical Report, National Bureau of Economic Research 2020.

_ , Steven M Ewer, and Todd H Wagner, "Returns to physician human capital: Evidence from patients randomized to physician teams," *Journal of health economics*, 2010, *29* (6), 866–882.

Kim, Scott and Petra Moser, "Women in Science. Lessons from the Baby Boom," *Econometrica*, 2025, *93* (5), 1521–1560.

Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante, "The Child Penalty Atlas," *National Bureau of Economic Research Working Paper Series*, August 2023, (31649).

_ , _ , and Jakob Egholt Søgaard, "Does biology drive child penalties? Evidence from biological and adoptive families," *American Economic Review: Insights*, 2021, *3* (2), 183–198.

_ , _ , and Jakob Egholt Søgaard, "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, October 2019, *11* (4), 181–209.

_ , _ , Johanna Posch, Andreas Steinhauer, and Josef Zweimüller, "Child penalties across countries: Evidence and explanations," in "AEA Papers and Proceedings," Vol. 109 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2019, pp. 122–126.

Lim, Wen Hui, Chloe Wong, Sneha Rajiv Jain, Cheng Han Ng, Chia Hui Tai, M Kamala Devi, Dujeepa D Samarasekera, Shridhar Ganpathi Iyer, and Choon Seng Chong, "The unspoken reality of gender bias in surgery: a qualitative systematic review," *PLoS One*, 2021, *16* (2), e0246420.

Orchard, Edwina R, Helena JV Rutherford, Avram J Holmes, and Sharna D Jamadar, "Matrescence: lifetime impact of motherhood on cognition and the brain," *Trends in Cognitive Sciences*, 2023, *27* (3), 302–316.

Oster, Emily, "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Paim, Jairnilson, Claudia Travassos, Celia Almeida, Ligia Bahia, and James Macinko, "The Brazilian health system: history, advances, and challenges," *The Lancet*, 2011, *377* (9779), 1778–1797.

Phillips, Susan P and Emily B Austin, "The feminization of medicine and population health," *Jama*, 2009, *301* (8), 863–864.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe, "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Sarsons, Heather, "Interpreting signals in the labor market: evidence from medical referrals," 2022, pp. 141–145.

Silver, David, "Haste or waste? Peer pressure and productivity in the emergency department," *The Review of Economic Studies*, 2021, *88* (3), 1385–1417.

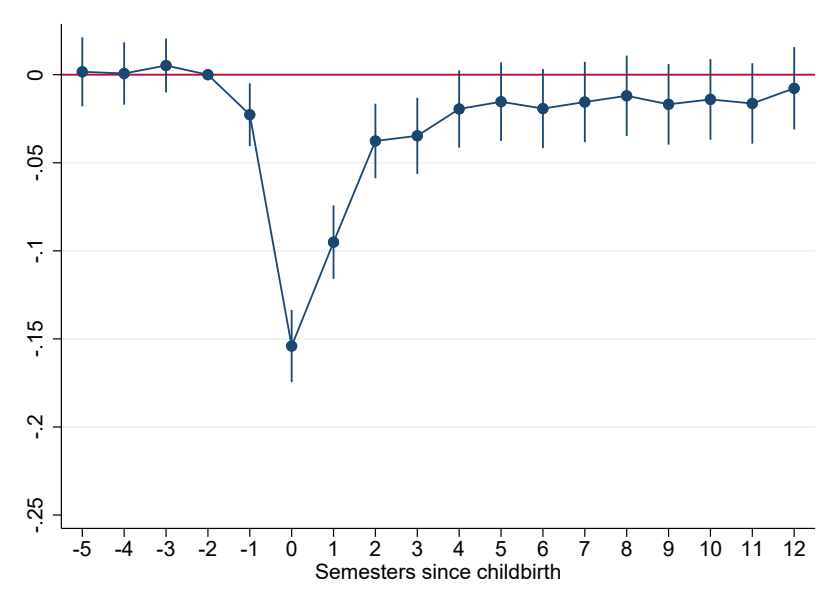
Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick, "Capitalists in the twenty-first century," *The Quarterly Journal of Economics*, 2019, *134* (4), 1675–1745.

Swain, Annette M, Michael W O'Hara, Kathleen R Starr, and Laura L Gorman, "A prospective study of sleep, mood, and cognitive function in postpartum and nonpostpartum women," *Obstetrics & Gynecology*, 1997, *90* (3), 381–386.

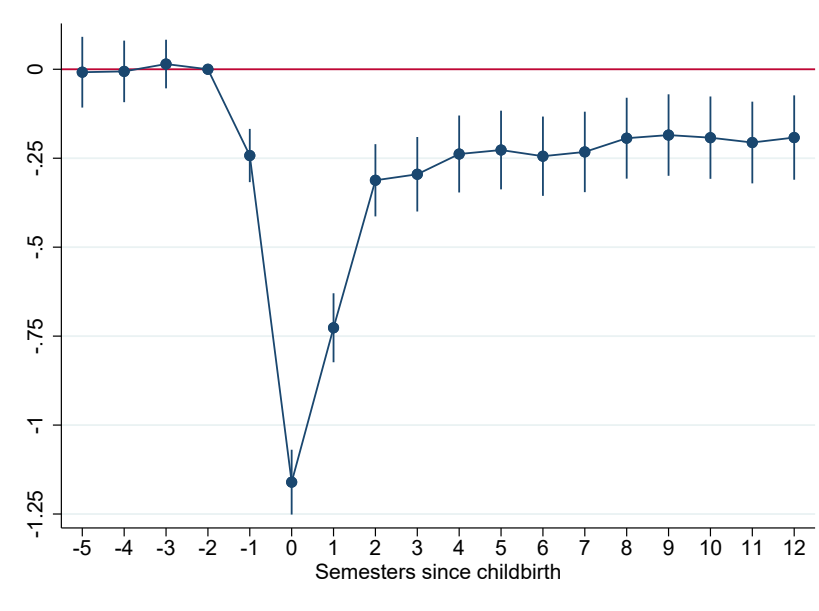
Trockel, Mickey T, Nikitha K Menon, Susannah G Rowe, Miriam T Stewart, Randall Smith, Ming Lu, Peter K Kim, Mariah A Quinn, Elizabeth Lawrence, Daniel Marchalik et al., "Assessment of physician sleep and wellness, burnout, and clinically significant medical errors," *JAMA network open*, 2020, *3* (12), e2028111.

Tsugawa, Yusuke, Anupam B Jena, Jose F Figueroa, E John Orav, Daniel M Blumenthal, and Ashish K Jha, "Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians," *JAMA internal medicine*, 2017, *177* (2), 206–213.

Figure 1: Effects of motherhood on payments for medical services



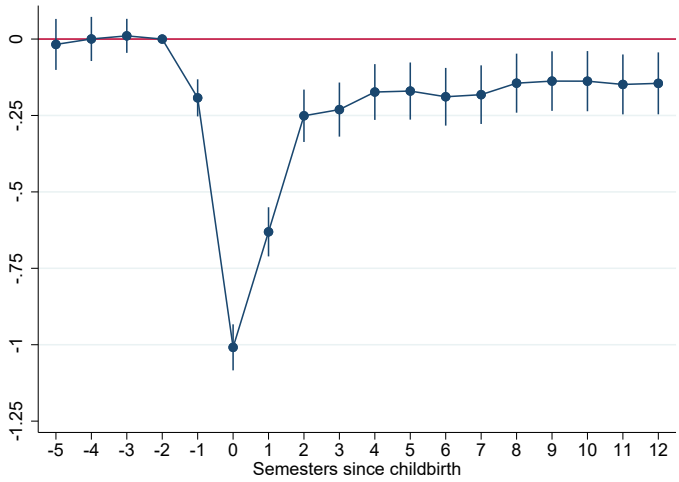
Panel A. Any payment



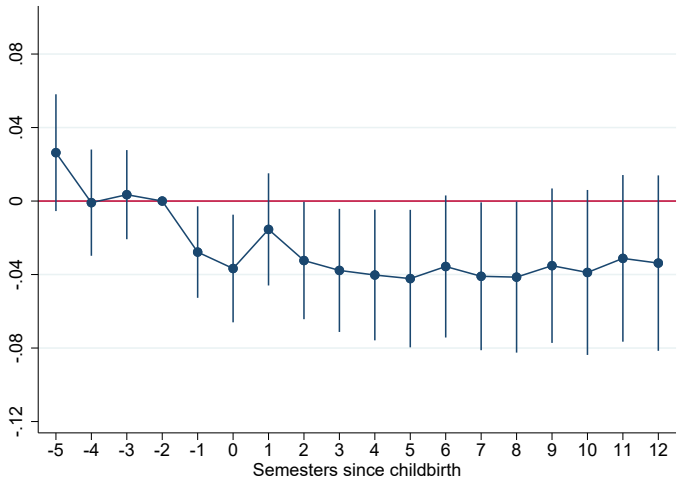
Panel B. Log of payments

Notes: This figure presents estimates based on the model (1) with respective 95 percent confidence intervals. The dependent variable in Panel A is a dummy that equals one if the physician receives any payment in a given semester. The dependent variable in panel B is the log of payments for medical services. To account for the fact that a few physician-semesters have no payments, we calibrate the extensive margin by substituting the log of zero payments with the log of the 1st percentile positive value of payments. Standard errors are clustered at the physician level

Figure 2: Effects of motherhood on volume of cases and payments per case, and decomposition of payments decline



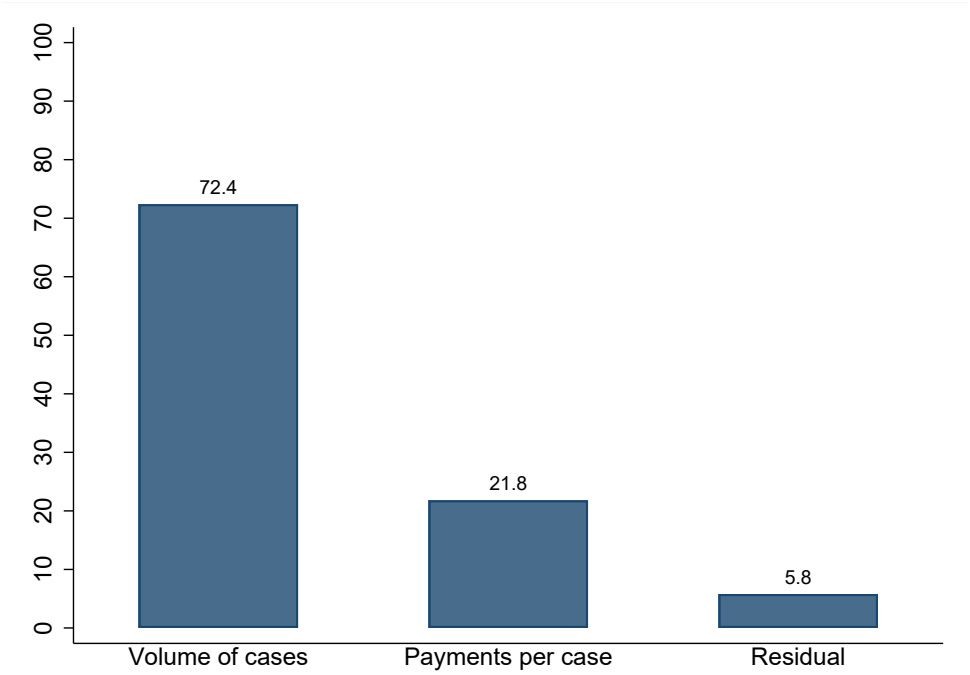
Panel A. Log of cases



Panel B. Log of payments per case

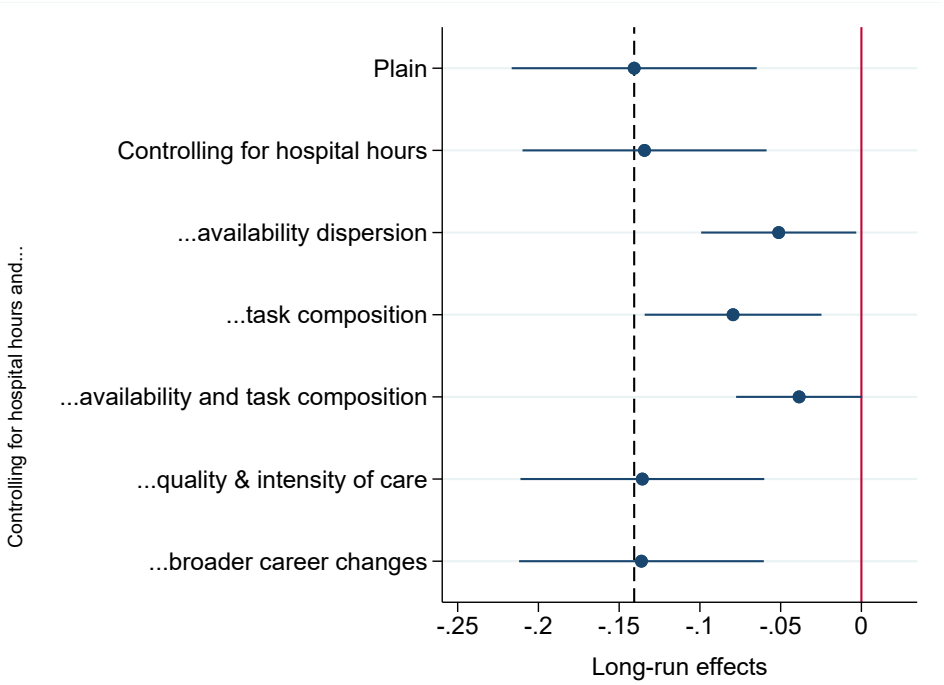
Notes: This figure presents estimates based on the model (1) with respective 95 percent confidence intervals. The dependent variable in Panel A is the log volume of cases. We calibrate the extensive margin by substituting the log of zero cases with the log of the 1st percentile positive value of cases. The dependent variable in Panel B is the log of payments per case. Standard errors are clustered at the physician level.

Figure 3: Decomposition of the long-run payments decline



Notes: This figure decomposes the long-run effects on log payments into reduced case volume, payments per case, and a residual, following the approach outlined in Section 4.3.

Figure 4: Long-run effects of motherhood on log of cases — controlling for mechanisms



Notes: This figure presents the long-run effects of motherhood on the log volume of cases, along with 95% confidence intervals, based on model (2) and analogous specifications that additionally control for different blocks of mechanisms discussed in Section 5.

Table 1: Effects of motherhood on risk-adjusted measures for quality and intensity of care

	(1)	(2)	(3)	(4)	(5)
	Mortality	30-day Readmission	Spendings	Procedures	Lenght of Stay
Short-run	-0.0007	0.0046**	-13.4140	-0.0418	0.0736
	(0.0011)	(0.0022)	(17.8999)	(0.0886)	(0.0660)
Medium-run	-0.0003	0.0016	-11.0518	0.0976	0.1212
	(0.0012)	(0.0026)	(21.4954)	(0.1116)	(0.0773)
Long-run	0.0000	-0.0008	18.1006	-0.0701	0.1623*
	(0.0014)	(0.0028)	(27.6987)	(0.1519)	(0.0879)
Pre-event mean	0.043	0.108	468.845	8.032	5.378
Observations	116,970	116,970	116,970	116,970	116,970

Notes. This table presents estimates based on the model (2). Columns (1) and (2) investigate, respectively, the mortality and 30-day readmission rates of a physician's patients. Columns (3), (4), and (5) investigate, respectively, the average spending, number of procedures, and length of stay per patient. Outcomes in all columns are risk-adjusted, using the two-step procedure described in Section 2.2. Standard errors are clustered at the physician level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Effects of motherhood on dispersion availability

	(1)	(2)	(3)	(4)
	# Hospitals	Hours per hospital	# Workdays	Hours per workday
Short-run	-0.0680***	0.4498**	-0.5540***	1.9207***
	(0.0175)	(0.1989)	(0.0377)	(0.2267)
Medium-run	-0.1093***	1.2676***	-0.1680***	0.5831**
	(0.0219)	(0.2351)	(0.0428)	(0.2448)
Long-run	-0.1565***	1.4310***	-0.1735***	0.5242*
	(0.0257)	(0.2661)	(0.0491)	(0.2758)
Pre-event mean	2.085	19.992	5.291	10.089
Observations	167,189	167,189	127,019	127,019

Notes. This table presents estimates based on the model (2). Column (1) investigates the number of hospitals a physician practices. Column (2) examines the number of work hours per affiliated hospital. Column (3) assesses the number of calendar days a physician practices. Column (4) investigates the number of work hours per workday. Standard errors are clustered at the physician level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Effects of motherhood on output composition within hospitals

	(1)	(2)	(3)	(4)	(5)	(6)
	1(Surgery)	1(Elective)	1(High-complexity)	# Diagnoses	% Supporting	% Non-clinical
Short-run	-0.0156***	-0.0330***	-0.0367***	-3.9871***	0.0177***	0.0048*
	(0.0044)	(0.0079)	(0.0062)	(0.2859)	(0.0058)	(0.0025)
Medium-run	-0.0099**	-0.0114	-0.0104	-0.8715**	0.0151**	0.0061**
	(0.0049)	(0.0089)	(0.0069)	(0.3530)	(0.0073)	(0.0026)
Long-run	-0.0094*	-0.0206**	-0.0194**	-0.7233*	0.0041	0.0075***
	(0.0055)	(0.0102)	(0.0081)	(0.4171)	(0.0097)	(0.0028)
Pre-event mean	0.383	0.363	0.193	14.418	0.537	0.091
Observations	116,970	116,970	116,970	116,970	116,970	167,184

Notes. This table presents estimates based on the model (2). Columns (1), (2), and (3) examine the probability that the physician performs surgeries, elective procedures, and high-complexity procedures, respectively. Column (4) examines the number of distinct diagnoses the physician treats. Column (5) reports the share of inpatient cases in which the physician acts as a collaborator rather than the case leader. Column (6) reports the share of within-hospital work hours that the physician devotes to non-clinical duties. Standard errors are clustered at the physician level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Effects of motherhood on broader career changes

	(1)	(2)	(3)	(4)
	New Job	New Occupation	% Hospital	% SUS
Short-run	-0.0393***	-0.0430***	-0.0063	0.0093***
	(0.0104)	(0.0090)	(0.0046)	(0.0034)
Medium-run	-0.0232**	-0.0473***	0.0037	0.0182***
	(0.0102)	(0.0095)	(0.0059)	(0.0048)
Long-run	-0.0055	-0.0400***	0.0127*	0.0237***
	(0.0098)	(0.0097)	(0.0069)	(0.0061)
Pre-event mean	0.000	0.000	0.747	0.888
Observations	180,557	180,557	175,733	175,733

Notes. This table presents estimates based on the model (2). Columns (1) and (2) examine, respectively, the probability that the physician moves toward a new health facility and medical occupation. Columns (3) and (4) investigate, respectively, the share of the physician's work hours devoted to hospitals and SUS patients. Standard errors are clustered at the physician level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix to "Motherhood Penalty in Labor Output: Evidence from Physicians"

Danyelle Branco
Sao Paulo School of Economics FGV

Bladimir Carrillo
Sao Paulo School of Economics FGV

Luiz Fontes
Institute of Education and Research - INSPER

Laisa Rachter
IDB

January 28, 2026

List of Figures 1

List of Tables 2

A Additional Tables and Figures 3

List of Figures

A.1 Timing of motherhood events 3
A.2 Long-run effects of motherhood on log of payments for medical services and volume of cases — alternative calibrations of the extensive margin 4
A.3 Effects of motherhood on log of volume of cases and cases per hospital hour — physicians with positive caseload 5
A.4 Effects of motherhood on log of payments, volume of cases, and payments per case — extended post-treatment window 6
A.5 Long-run effects of motherhood on log of payments, volume of cases, and payments per case — excluding prior childbirths 7

Contact information: Branco: Sao Paulo School of Economics-FGV, Sao Paulo, SP, Brazil, 01332-000 (e-mail: danyelle.branco@fgv.br). Carrillo: Sao Paulo School of Economics-FGV, Sao Paulo, SP, Brazil, 01332-000 (e-mail: bladimir.carrillo@fgv.br). Fontes: Insper - Institute of Education and Research, Sao Paulo, SP, Brazil, 04546-042 (e-mail: luiz.fontes@insper.edu.br). Rachter: Inter-American Development Bank, Brasilia, Brazil, 70800-400 (e-mail: laisarachter@gmail.com).

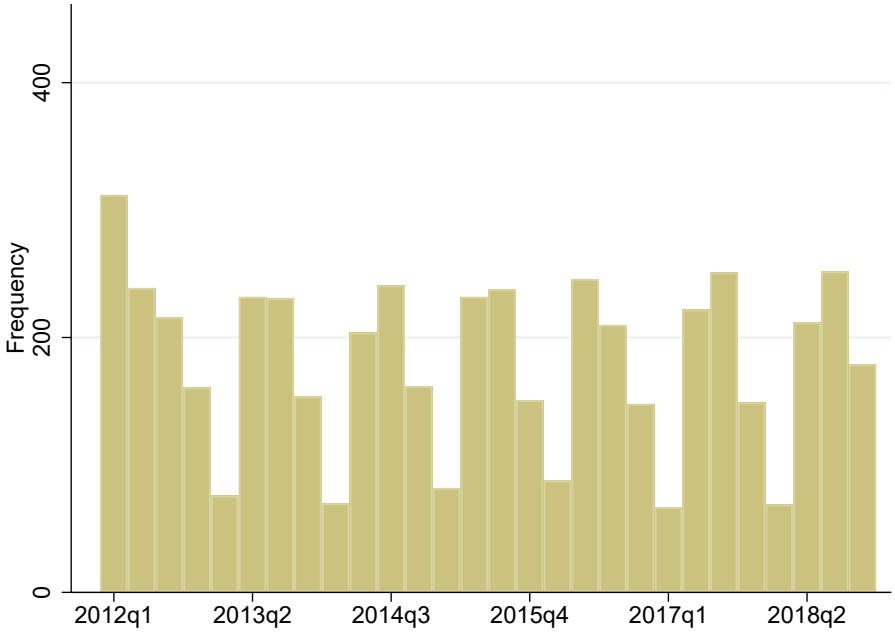
A.6 Inference for the long-run effects of motherhood on log payments, log volume of cases, and log payments per case across different clustering choices 8

List of Tables

A.1 Balancing Table 9
A.2 Effects of motherhood on log of work hours and payments per work hour 10
A.3 Test of joint significance of pre-trends — main outcomes 11
A.4 Effects of motherhood on log of payments, volume of cases, and payments per case — robustness to matching-strata-specific time effects 12
A.5 Test of joint significance of pre-trends — mechanisms 13
A.6 Effects of motherhood on the share of each calendar day within the physician's overall workload 14

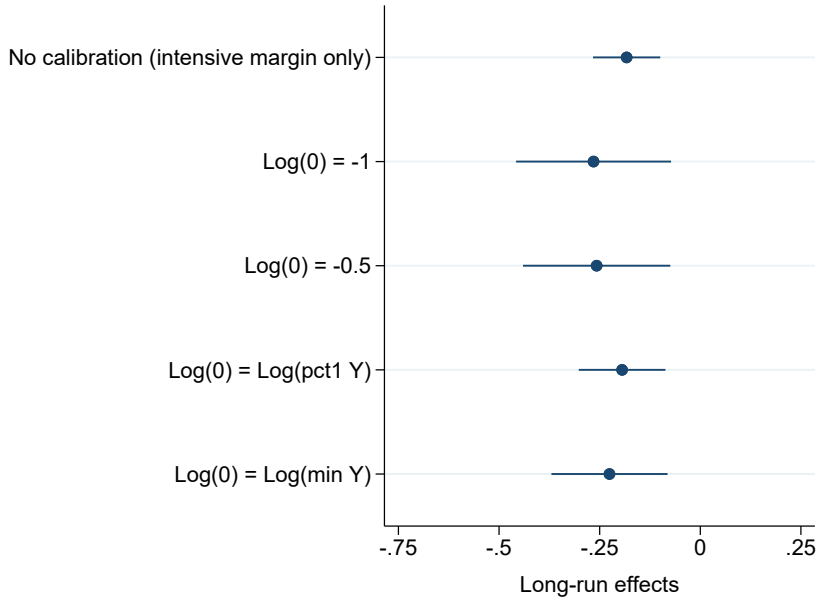
A Additional Tables and Figures

Figure A.1: Timing of motherhood events

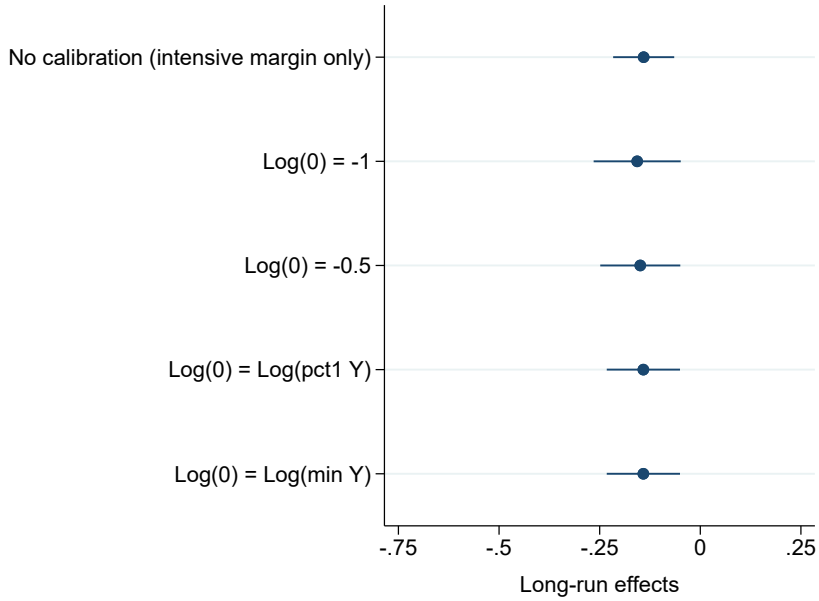


Notes: This figure plots the distribution of motherhood events over time.

Figure A.2: Long-run effects of motherhood on log of payments for medical services and volume of cases — alternative calibrations of the extensive margin



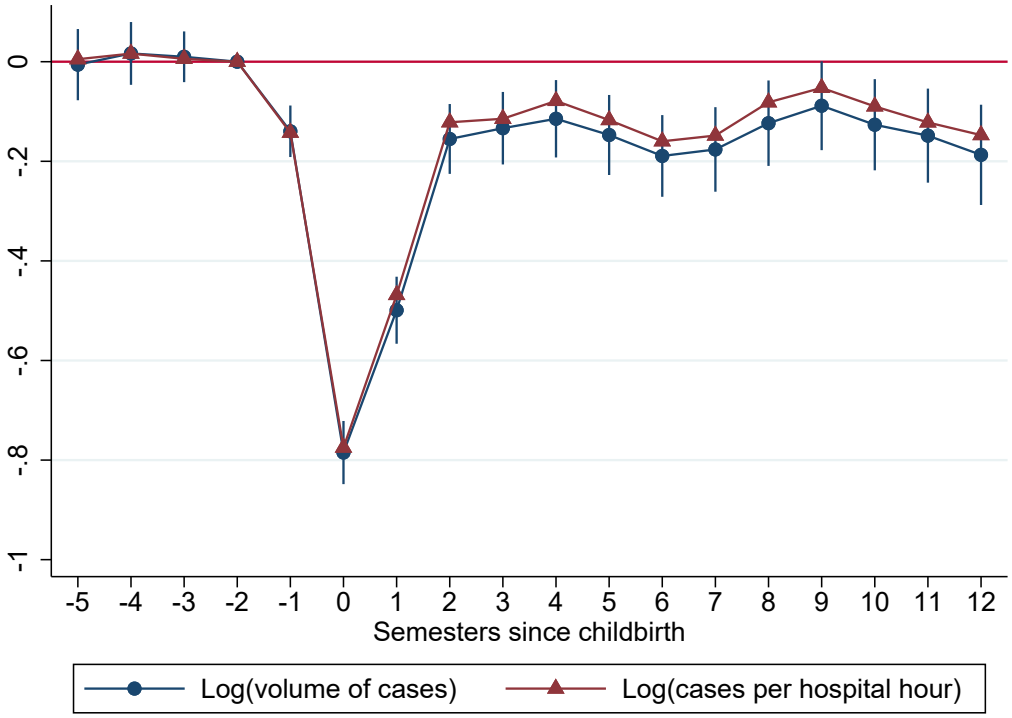
Panel A. Log of payments



Panel B. Log of cases

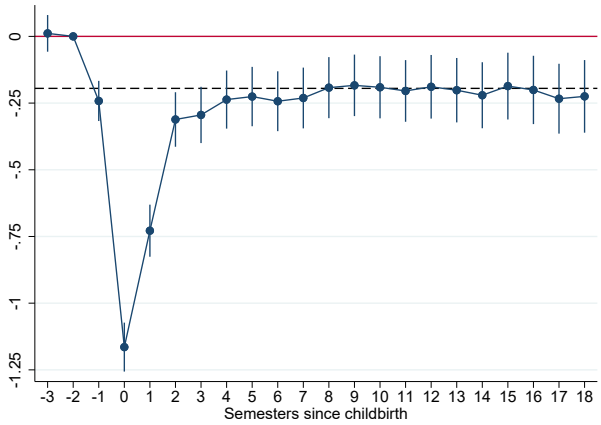
Notes: This figure presents the long-run effects of motherhood on the log payments and the log volume of cases. Estimates are obtained from model (2) and shown with 95 percent confidence intervals. We present the results without calibrating the extensive margin (i.e., using only physicians with positive payments or cases), and across different strategies for calibration following Chen and Roth (2024). Standard errors are clustered at the physician level.

Figure A.3: Effects of motherhood on log of volume of cases and cases per hospital hour — physicians with positive caseload

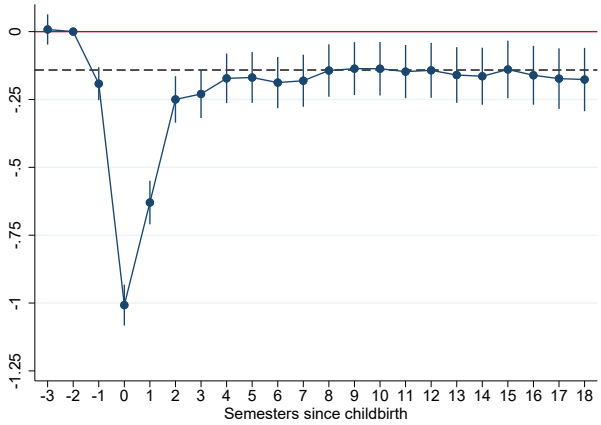


Notes: This figure presents estimates based on the model (1) with respective 95 percent confidence intervals. The dependent variables are the log volume of cases and the log volume of cases per hospital hour. Standard errors are clustered at the physician level.

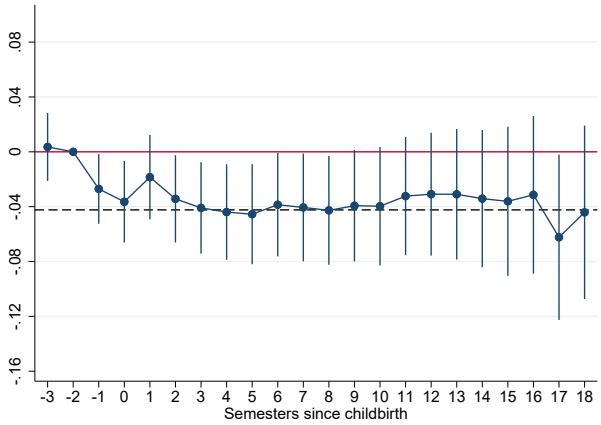
Figure A.4: Effects of motherhood on log of payments, volume of cases, and payments per case — extended post-treatment window



Panel A. Log of payments



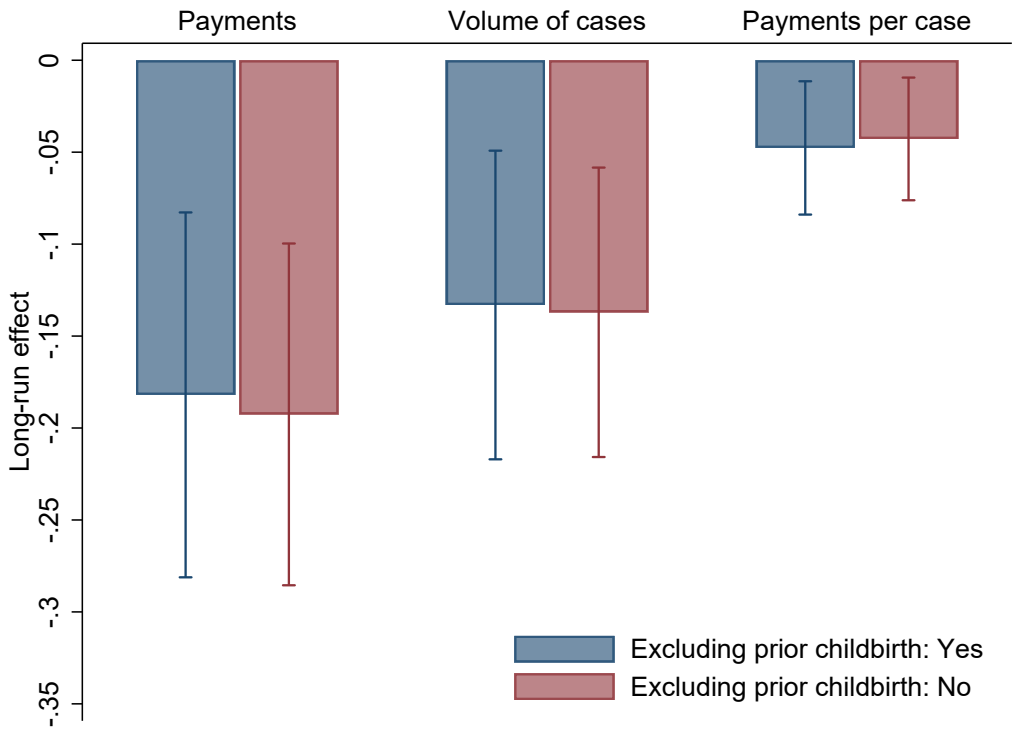
Panel B. Log of cases



Panel C. Log of payments per case

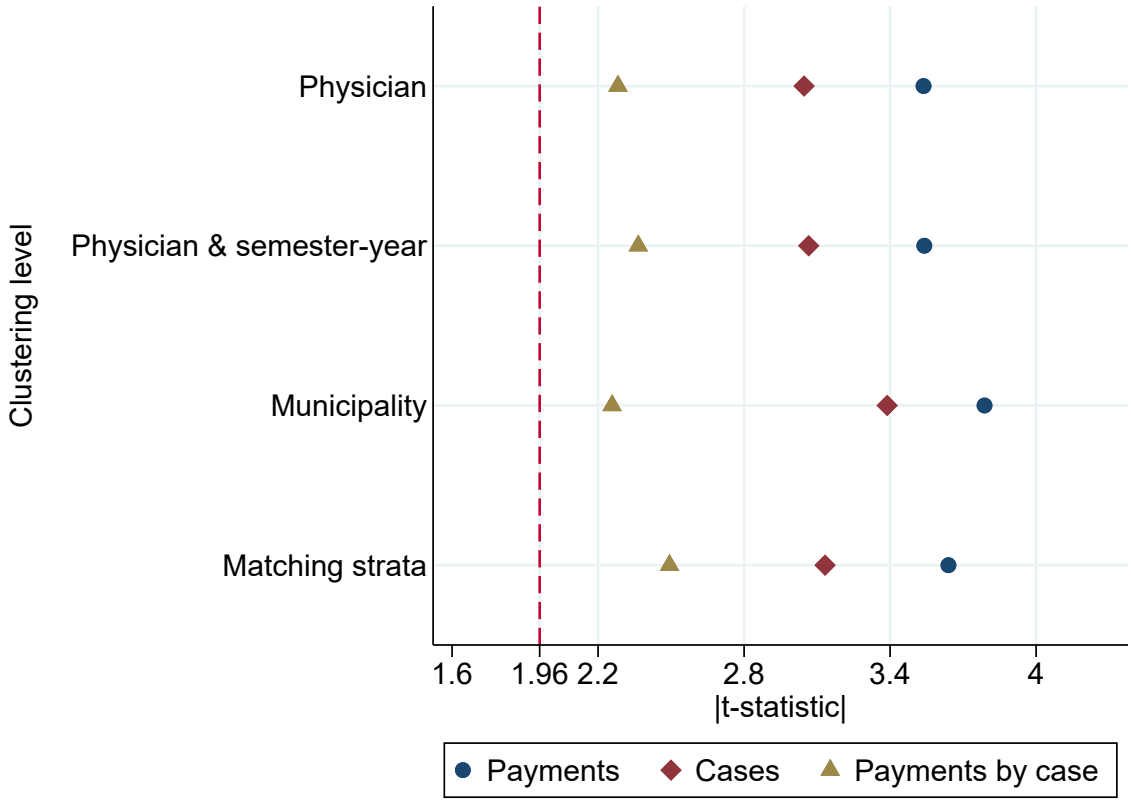
Notes: This figure presents estimates based on the model (1) with respective 95 percent confidence intervals in an extended sample following female physicians up to 18 semesters after childbirth. The dashed line represents the long-run effects of motherhood estimated based on the model (2) using our baseline sample. The dependent variable in panel A is the log of payments for medical services. The dependent variable in panel B is the log volume of cases. To account for the fact that a few physician-semesters have no payments (cases), we calibrate the extensive margin by substituting the log of zero payments (cases) with the log of the 1st percentile positive value of payments (cases). The dependent variable in panel C is the log payments per case. Standard errors are clustered at the physician level.

Figure A.5: Long-run effects of motherhood on log of payments, volume of cases, and payments per case — excluding prior childbirths



Notes: This figure presents the long-run effects of motherhood on log payments, the log volume of cases, and log payments per case. Estimates are obtained from model (2) and shown with 95 percent confidence intervals. Results are reported for the baseline sample and for a secondary sample that excludes female physicians who experienced a maternity leave spell prior to the study period (2008–2024). Standard errors are clustered at the physician level.

Figure A.6: Inference for the long-run effects of motherhood on log payments, log volume of cases, and log payments per case across different clustering choices



Notes: This figure presents the absolute value of the *t*-statistics for the long-run effects of motherhood on log payments, the log volume of cases, and log payments per case. Estimates are obtained from model (2) and are constructed under different clustering levels: physician, physician and semester-year, municipality of practice, and matching strata.

Table A.1: Balancing Table

	(1) Mothers	(2) Matched Nonmothers	(3) All Nonmothers
Years of medical experience	8.078	8.059	16.412
	(3.780)	(3.742)	(11.185)
Top 25 Medical School	0.322	0.321	0.363
	(0.467)	(0.467)	(0.481)
Advanced specialization	0.757	0.746	0.467
	(0.429)	(0.435)	(0.499)
Surgery practice	0.381	0.385	0.357
	(0.486)	(0.487)	(0.479)
Physician license: North	0.038	0.037	0.048
	(0.191)	(0.189)	(0.214)
Physician license: Northeast	0.155	0.151	0.204
	(0.362)	(0.358)	(0.403)
Physician license: Southeast	0.549	0.555	0.552
	(0.498)	(0.497)	(0.497)
Physician license: Centerwest	0.159	0.159	0.125
	(0.365)	(0.365)	(0.331)
Physician license: South	0.099	0.099	0.070
	(0.299)	(0.298)	(0.256)
Payments	11444.521	11240.968	13068.992
	(37534.008)	(26462.741)	(37793.595)
Volume of cases	91.013	91.952	106.451
	(147.006)	(153.574)	(218.965)
Payments per case	145.676	141.390	145.304
	(378.330)	(212.995)	(281.124)
Weekly work hours	52.509	51.473	49.750
	(25.832)	(26.499)	(25.533)
Hospitals	2.054	2.044	1.683
	(1.112)	(1.153)	(1.092)
Calendar workdays	5.338	5.245	5.198
	(2.066)	(2.105)	(2.164)
% Work hours within hospitals	0.735	0.758	0.665
	(0.293)	(0.290)	(0.345)
% Work hours within SUS	0.888	0.888	0.885
	(0.184)	(0.191)	(0.251)
% Non-clinical duties	0.088	0.094	0.078
	(0.213)	(0.224)	(0.204)
Perform elective procedures	0.350	0.376	0.350
	(0.477)	(0.484)	(0.477)
Perform high-complexity procedures	0.179	0.207	0.184
	(0.384)	(0.405)	(0.387)
% Leadership of cases	0.578	0.574	0.582
	(0.310)	(0.304)	(0.310)
Mortality rate	0.061	0.065	0.063
	(0.153)	(0.153)	(0.158)
Readmission rate	0.105	0.111	0.102
	(0.169)	(0.178)	(0.169)
Lenght of stay	7.185	6.914	7.126
	(8.334)	(7.363)	(7.836)
Number of procedures	9.488	9.153	8.738
	(7.612)	(7.540)	(7.679)
Spendings	931.478	817.452	842.135
	(2556.040)	(2209.241)	(2286.863)
Number of unique physicians	4815	4815	56692

Notes. This table shows summary statistics for physicians in our data. In columns (1) and (2), variables are measured four quarters before the childbirth date (placebo date for matched nonmothers). In column (3), variables are measured using the entire data, but each period is weighted by the share of physicians who experienced a childbirth event four quarters later. Standard deviations in parentheses.

Table A.2: Effects of motherhood on log of work hours and payments per work hour

	(1) Log(Work hours)	(2) Log(Payments per work hour)
Short-run	0.0044	-0.7410***
	(0.0120)	(0.0463)
Medium-run	0.0081	-0.2412***
	(0.0147)	(0.0532)
Long-run	-0.0078	-0.1642***
	(0.0162)	(0.0572)
Pre-event mean	6.975	0.571
Observations	180,557	175,733

Notes. This table presents estimates based on the model (2). Column (1) examines the log of work hours. Column (2) investigates the log of payments per work hour. Standard errors are clustered at the physician level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Test of joint significance of pre-trends — main outcomes

	Pre-trend window: 5		Pre-trend window: 8	
	F-statistic	p-value	F-statistic	p-value
	(1)	(2)	(3)	(4)
Log(payments)	0.148	0.931	0.372	0.897
Log(volume of cases)	0.195	0.900	0.496	0.812
Log(payments per case)	1.449	0.226	1.022	0.408

Notes. This table reports the *F*-statistic and *p*-value from *F*-tests under the null hypothesis that all pre-treatment coefficients in our baseline specifications are jointly zero. We estimate pre-treatment coefficients using model (1). Columns (1) and (2) consider our baseline pre-treatment window of 5 periods. Columns (3) and (4) extend the pre-treatment window to include 8 periods.

Table A.4: Effects of motherhood on log of payments, volume of cases, and payments per case — robustness to matching-strata-specific time effects

	Log(payments)		Log(volume of cases)		Log(payments per case)	
	(1)	(2)	(3)	(4)	(5)	(6)
Short-run	-0.7335***	-0.7412***	-0.6290***	-0.6326***	-0.0345***	-0.0466***
	(0.0450)	(0.0469)	(0.0376)	(0.0392)	(0.0123)	(0.0137)
Medium-run	-0.2479***	-0.2521***	-0.1879***	-0.1901***	-0.0460***	-0.0564***
	(0.0515)	(0.0533)	(0.0433)	(0.0448)	(0.0152)	(0.0168)
Long-run	-0.1943***	-0.1926***	-0.1414***	-0.1371***	-0.0424**	-0.0428**
	(0.0549)	(0.0567)	(0.0464)	(0.0480)	(0.0186)	(0.0203)
Time-by-cohort FE	Yes		Yes		Yes	
Time-by-matching-strata FE		Yes		Yes		Yes
R2	0.583	0.700	0.575	0.692	0.720	0.811
Observations	180,557	172,032	180,557	172,032	116,940	98,085

Notes. The odd column numbers of this table present estimates based on the model (2). The even column numbers report the effects from a similar spefication, where we substitute the time$\times$ cohort fixed-effects by time$\times$matching-strata fixed-effects. Columns (1)-(2) investigate the log payments for medical services. Columns (3)-(4) examine the log volume of cases. Columns (5)-(6) turn to the log of payments per case. Standard errors are clustered at the physician level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Test of joint significance of pre-trends — mechanisms

	Pre-trend window: 5		Pre-trend window: 8	
	F-statistic	p-value	F-statistic	p-value
	(1)	(2)	(3)	(4)
Mortality	0.656	0.579	0.770	0.593
Readmission	0.158	0.924	0.750	0.610
Spending	0.282	0.839	0.950	0.457
Procedures	0.761	0.516	0.701	0.649
Length of stay	0.058	0.982	0.631	0.706
Hospitals	1.738	0.157	1.774	0.100
Workdays	0.688	0.559	0.491	0.815
1(Surgery)	0.054	0.983	0.950	0.458
1(Elective)	0.903	0.439	0.968	0.445
1(High-complexity)	0.604	0.613	0.741	0.617
Diagnoses	0.147	0.932	0.125	0.993
% Non-clinical	0.363	0.780	0.735	0.622
% Supporting	0.932	0.424	0.661	0.682
New Job	0.062	0.980	0.481	0.823
New Occupation	0.224	0.880	0.124	0.994
% Hospital	2.274	0.078	1.454	0.190
% SUS	0.111	0.954	0.386	0.889

Notes. This table reports the *F*-statistic and *p*-value from *F*-tests under the null hypothesis that all pre-treatment coefficients in our baseline specifications are jointly zero. We estimate pre-treatment coefficients using model (1). Columns (1) and (2) consider our baseline pre-treatment window of 5 periods. Columns (3) and (4) extend the pre-treatment window to include 8 periods.

Table A.6: Effects of motherhood on the share of each calendar day within the physician's overall workload

	(1) % Monday	(2) % Tuesday	(3) % Wednesday	(4) % Thursday	(5) % Friday	(6) % Saturday	(7) % Sunday	(8) % Weekend
Short-run	-0.0019	0.0015	0.0002	0.0009	0.0029*	-0.0013	-0.0022	-0.0035*
	(0.0018)	(0.0018)	(0.0017)	(0.0017)	(0.0016)	(0.0012)	(0.0013)	(0.0020)
Medium-run	-0.0010	0.0027	0.0003	-0.0020	0.0040**	-0.0009	-0.0031**	-0.0040*
	(0.0021)	(0.0021)	(0.0021)	(0.0020)	(0.0018)	(0.0014)	(0.0016)	(0.0024)
Long-run	0.0051**	0.0042*	-0.0010	-0.0032	0.0011	-0.0027*	-0.0034*	-0.0061**
	(0.0024)	(0.0024)	(0.0024)	(0.0023)	(0.0021)	(0.0016)	(0.0018)	(0.0027)
Pre-event mean	0.161	0.159	0.157	0.155	0.146	0.109	0.113	0.229
Observations	127,652	127,652	127,652	127,652	127,652	127,652	127,652	127,652

Notes. This table presents estimates based on the model (2). Columns (1)-(7) report the effects on how much each of the respective calendar days represents within the calendar days the physician practices. Column (8) reports similar effects aggregating Saturday and Sunday. Standard errors are clustered at the physician level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.