Household Penalty: Gendered Costs of Spousal Infectious Diseases on Labor and Health*

Xinming Du† Krzysztof Zaremba‡

February 2025

Abstract

This paper examines the gendered impact of spousal health shocks on labor and health outcomes. Using Mexican labor surveys and a difference-in-differences approach, we find women's labor supply decreases by 15% after a partner's health shock, compared to a 10% reduction for men. A significant part of this is driven by women's higher probability of household infection. Analysis of U.S. insurance claims shows a partner's infection increases infection risk by 1.2 percentage points for men and 2.2 for women. Household specialization underpins these effects: as women's income share rises, their penalty decreases while men's increases.

JEL Classifications: I14, J22, O12, O15

Keywords: Household disease transmission, gender, home production, labor supply

*We are grateful to Sandra Aguilar-Gómez, Douglas Almond, Belinda Archibong, Joséphine Gantois, Wolfram Schlenker, Jeffrey Shrader, Yulya Truskinovsky, Emma Zai and participants at MEA, ASHEcon, Workshop on Healthy Ageing and Adult Caregiving (LSE/UPenn), and Columbia University for helpful comments and suggestions. The data, technology, and services used in the generation of these research findings were generously supplied pro bono by the COVID Research Database partners, who are acknowledged at https://covid19researchdatabase.org/, K. Zaremba acknowledges support from the Asociacion Mexicana de Cultura.

†Department of Economics, National University of Singapore and Center for Environmental Economics and Policy, Columbia University, Email: xd2197@columbia.edu

‡Instituto Tecnológico Autónomo de México, E-mail: zaremba@itam.mx

1 Introduction

Within households, agents frequently engage in interactions that give rise to externalities, many of which bear significant implications for overall welfare. A salient example of such an externality is the transmission of health shocks. Here, a negative health event affecting one agent can precipitate a series of repercussions on the health, productivity, and labor supply decisions of other household members (Daysal et al., 2021). This phenomenon offers a unique perspective for analyzing intra-household resource allocation, especially under the lens of infectious disease transmission and its consequences for health outcomes.

This framework of health-related externalities is intricately linked to the gender dynamics within households, particularly emphasizing the role of women as primary caregivers. Women are inherently subject to distinct health challenges, influenced by biological factors, economic barriers, and societal norms. In this context, their often-assumed caregiving roles—rooted in social expectations and economic constraints—further exacerbate their vulnerability to health shocks.

Building on these themes, our paper explores the interaction between intra-household health shock externalities and gender-based health and economic disparities. We investigate whether female agents, vis-a-vis their male counterparts, incur a disproportionate health and economic cost when another agent within the household experiences a health shock. By analyzing the responses of female health conditions and labor supply decisions to these shocks, we aim to shed light on the economic mechanisms driving these outcomes.

We introduce a concept of "household penalty", defined as the spillover effect of family members' health shocks on other members' health conditions and labor outcomes. Focusing on infectious disease transmission, we build a theoretical framework of the health production model as an extension of Grossman (2000). We assume an individual's health depends on their previous health stock and other family members' investment, which is a household optimization problem. Due to the negative impact of providing care to an infected family member on the caregiver's health, our model derives optimal caregiving time allocated by male and female members. We calculate gender-specific caregiving time and equilibrium health outcomes. Women face extra penalties if they have comparative advantages in providing care or have lower incomes than male members.

We test household penalties in two empirical settings.

First, using data from Mexico, we show that a partner's illness negatively impacts individual health and labor supply, with this externality being particularly pronounced for women. Specifically, we use data from the Mexican Labor Survey (ENOE) that encompasses approximately 23,000 households and has a 5-period rotating panel structure. It is representative of the population and covers both formal and informal employment. Our empirical strategy is a difference-in-difference with the interaction between the female gender and having a sick partner serving as the estimator for the extent of the female penalty. We also use event-study to assess the parallel trends assumption. We primarily anchor our analysis on two dependent variables: the likelihood of an individual abstaining from work in the preceding week due to sickness or personal reasons, and the cumulative working hours during the same period. The results show that a male partner's sickness triggers an 13.2-percentage-point increase in the probability of their female counterpart also being sick, a stark contrast to the 8.1 percentage point rise for males when the scenario is inversed. Regarding labor hours, sick male partners precipitate a 15% (4.7-hour) reduction in women' work, while men curtail their hours by 10% (4.2) when faced with a sick female partner. Female penalty in terms of hours is considerably larger (6.3 hour vs 4.3 hour) if we focus on full time workers. These results clearly demonstrate a gender-asymmetric externality from a partner's health shock. However, the ENOE data alone do not fully reveal the underlying reasons for this externality

To more precisely investigate the channel of infection transmission posited by our model, we turn to US insurance claims data, which offers detailed, time-stamped information on disease transmission within households, matched with members' characteristics. We quantify the higher probability of getting infected under the condition of an earlier infection of a family member, i.e., the household penalty, and we reveal how this penalty differs by gender. Under a difference-in-difference design, we find the probability of infection is 1.2 percentage points higher for male members and 2.2 percentage points higher for female members due to household contagion, namely 12.4% and 22.7% higher relative to the average infection rate. The results are robust after controlling for household size, household income, home value, and the number of children and adults in the household. Disentangling the effect by member's gender, we find no significant evidence of male-to-male transmission, while male-to-female and female-to-male transmission increases the exposed individual's infection rate by 2.8 and 1.5 percentage points. Extra household penalty for women is more striking when the first infected patient is middle-aged and in areas with a striking gender gap in income and in home production intensity.

Moreover, the results from the Mexican data confirm the model's predictions regarding the role of income composition within the household. Transitioning from a male-only earner household to a female-only earner household eradicates the added female penalty in falling sick and in working hours. This observation aligns with the notion that household specialization acts as a key mechanism in driving the gender disparity in household penalty.

Another mechanism lies in the uneven burden of home production faced by female members, such as in housework and childcare. We use time-use survey data and find particularly pronounced gender disparities in areas with more substantial gender gaps in home production time. Women tend to shoulder a greater share of the housework burden compared to their partners. In the context of infectious diseases, women' significant caregiving contributions may be implicitly expected, necessitating their provision of care to infected patients.

Our heterogeneity analysis also demonstrates that the presence of other potential caregivers can insulate women from the household penalty. When an elderly female, potentially serving as an alternate caregiver, is present in the household, the additional female penalty is virtually erased. Presence of an elderly male does not reduce the additional female penalty. Moreover, we show that most of the gender gap in the response is driven by workers in the formal sector, while the gap is smaller in the informal sector.

Our paper provides a novel empirical investigation into gender inequality in household health spillovers. We examine how partners' transmittable infections affect the health and labor supply decisions of other household members. Previous epidemiological studies have primarily focused on health spillovers, highlighting the disproportionate infection rates among women due to household transmission. For instance, Nkangu et al. (2017) reported a higher incidence of Ebola cases among women compared to men, despite the absence of a biological gender vulnerability gap. They suggested that women's primary role as caregivers in households might be a contributing factor but did not provide numerical evidence to support this claim. Similar assertions can be found in the works of WHO (2007) and Skrip et al. (2017), which emphasize gendered roles in domestic work and the varying infection risks associated with these roles within households. Close to our study is Daysal et al. (2021), which focuses on the disease transmission across young children within a family and the implications for their future outcomes, and Arrieta and Li (2023), which analyses family members' employment responses following an ED visit in the US context. Complementing this work, our paper concentrates on the household transmission of infections among adult members driving the gender gap in health and labor supply in both developed and developing countries.

This study extends the current understanding of the determinants of female labor supply by elucidating the disproportionate impact of partners' health shocks on women's labor decisions. Existing papers in this field primarily emphasize factors such as childbirth and childcare (Kleven et al. (2019); Aguilar-Gomez et al. (2019)). However, the existing body of literature concerning caregiving to adults and its implications on the labor market predominantly addresses elderly care (Mommaerts and Truskinovsky, 2020, 2023) or care for adults with serious or permanent disabilities (Fadlon and Nielsen, 2019, 2021; Lee, 2020; Brito and Contrera, 2023). An exception in this context is the study by Maestas et al. (2023), which investigates the long-term labor market effects of caregiving on different age groups. These studies, however, are primarily confined to developed countries or focus only on formal employment responses. We add to this literature by analyzing responses to common, transient infectious diseases among working-age adults in both formal and informal sectors. Our study reveals substantial asymmetries in the health and labor outcomes of female versus male caregivers following the health shock.

Moreover, we complement previous literature by looking at a developing country (Mexico). In developing countries like Mexico, characterized by a substantial informal employment sector and pronounced gender norms, the repercussions of household health shocks on the labor supply are potentially more pronounced. By incorporating a representative sample including both formal and informal employment, our paper offers a comprehensive perspective on how household health shocks can disproportionately affect health and labor supply outcomes, thereby contributing to gender disparities.

Empirics aside, our paper is the first to apply infectious disease to the Grossman model of health production. Infectious disease provides a novel setting where initial health shocks affect exposed family members' health through two channels: the direct effect of being infected on health conditions, and the indirect effect through the income channel due to decreased time on labor supply. In contrast, existing literature on health production mostly focuses on non-infectious initial shocks that affect others' health through the indirect income channel (e.g., Yi et al., 2015; Barreca et al., 2016; Heckman et al., 2018). Additionally, we separately discuss the equilibrium across the severity of infectious disease. Our main model focuses on the optimal care-providing time and health conditions when the infectious disease does not pose an excessively high risk and the provision of care yields sufficient returns. In Section S1, we explore an extreme scenario where the infectious disease is highly transmittable, the cost of care provision is high, the optimal caregiving amount is zero, and external interventions are needed to enhance public health. This has real-world implications like pandemics and vaccination investment.

From the policy perspective, our paper highlights the pressing need for intensified efforts in health support and infectious disease prevention. The consequences of health burdens are not confined solely to an individual's health status but rather ripple through to impact family members' health and labor outcomes. Consequently, our paper underscores the critical importance of prioritizing preventive measures to mitigate the far-reaching socioeconomic effects of health crises.

Furthermore, this paper underscores the need to implement supportive policies aimed at benefiting female family members. Our findings show that when partners become ill, women bear a disproportionate burden, grappling with heightened health-related challenges, caregiving responsibilities, and exacerbated adverse labor outcomes. To rectify this inequity, it is important to undertake proactive measures, including the provision of paid leave options for female caregivers or care subsidies.

The rest of the paper is structured as follows. Section 2 describes the background and related literature. Section 3 shows a stylized model that motivates our empirical study. Section 4 describes the data. Section 6 shows the empirical strategy, main results, and additional results using survey data from Mexico. Section 5 provides additional results using the US insurance claim data. Section 7 concludes.

2 Background

2.1 Related literature: household disease transmission

Intra-household transmission accounts for a great proportion of the spread of infectious diseases such as Ebola (Glynn et al., 2018) and COVID (Chan et al., 2020). Some epidemiological studies quantify the transmission rate among family members (e.g. Curmei et al., 2020; Esteve et al., 2020; Li et al., 2020). They confirm it is crucial to take household transmission into account so as to control the disease spread. Moreover, Daysal et al. (2021) demonstrate that within-household transmission to children can have critical long-term consequences in terms of human capital formation and earnings. Despite the importance of transmission within a family, little is known about the unequal exposure of family caregivers. As the burden of infectious disease has fallen since 1890, Goldin and Lleras-Muney (2018) find women benefit more greatly from this public health improvement. They mention one possible reason is that female children had a greater role in taking care of sick family members, whereas the boys were out of the house more, possibly at work.

Gender, together with other covariates, is shown to be associated with different exposure, infection, attitude, and response towards infectious diseases. Lewandowski et al. (2020) suggest women are more likely to be exposed to contagion due to their sectoral segregation into occupations that require more interpersonal interactions. They show that gender is a more important factor in workers' exposure than education or age. Similar results are found by Chernoff and Warman (2020). They show that mid-educated women are at the highest risk of infection among automatable jobs, mainly because they work in healthcare, office and administrative support and protective service occupations. When it comes to attitude and avoidance behaviors, Galasso et al. (2020) find women are more likely to see the pandemic as a very serious health problem, to agree with restraining public policy measures adopted in response to it, and to comply with them. Also, Papageorge et al. (2020) finds women are more likely to engage in self-protective behaviors and believe in the effectiveness of social distancing than males. These studies suggest that there is a natural gender difference in infection rate regardless of household transmission.

Infectious disease contributes to gender inequality. Earlier pandemics have been shown to increase gender gaps in schooling and education attainment (Archibong and Annan, 2017, 2020). Earlier pandemic affects women' performance in the workplace and burden in housework. Infectious disease increases the demand for home childcare and women have been doing the great share (Sevilla and Smith, 2020). Mothers spend substantially longer on housework and sacrifice a larger fraction of their paid work hours than their partners (Andrew et al., 2020). As a result, women are more likely to lose jobs than men due to infectious disease and those still employed are more likely to work from home with disruptions and distractions (Farre et al., 2020; Papanikolaou and Schmidt, 2020).

Another aspect of gender inequality comes from mental health and utility of working in the presence of others. Gimenez-Nadal et al. (2020) find women benefit more from working with others in their daily activities than do men. As a result, solo work due to infectious diseases has more negative impacts on women's well-being. Survey evidence among prolific participants in the UK shows women are more concerned about getting and spreading the virus, perceive the virus as more prevalent and lethal, are more likely to expect a new lockdown or outbreak wave, and are more pessimistic about the current and future economy (Oreffice and Quintana-Domeque, 2020). Furthermore, it is widely discussed that pandemic and related lockdowns increase domestic violence (e.g., Ravindran and Shah, 2020; Leslie and Wilson, 2020), which further worsens gender inequality.

2.2 Related literature: labor supply and household caregiving

The interplay between household caregiving and employment has been reviewed by Lilly et al. (2007) and Bauer and Sousa-Poza (2015). The impact of informal caregiving on employment remains a topic of debate within the literature with mixed results. Brito and Contrera (2023) shows that adult daughters' employment decreases by 3% following their elderly parent's cancer diagnosis, while sons' employment remains unaffected. Maestas et al. (2023) find similar magnitudes for female caregivers under 50, which lasts for about 5 years, while men experience a decline in employment before the caregiving episode, and the recovery takes considerably longer. Fadlon and Nielsen (2021) use Danish administrative data to demonstrate that non-fatal, severe shocks to a spouse's health do not affect employment responses, as the formal insurance fills the missing income. Lee (2020) find no impact on female weekly working hours following husband's disability, but a 2-3 hour's increase in time spent on providing care. Coile et al. (2022) shows that negative impact on female employment can be mitigated by access to paid family leave. Arrieta and Li (2023) reveals an increase in female working hours by 0.5% following a family member's visit to the ED, while no impact on men, consistently with the insurance mechanism dominating the caregiving mechanism. They also show that the increase in labor supply is less pronounced for conditions requiring more caregiving. It's important to note that many of these dynamics, including formal insurance and paid family leave, are predominantly characteristic of affluent, developed nations. Consequently, the nature of these problems and their responses are likely to vary substantially in different socio-economic contexts. More substantial effects have been observed in cases of very intense caregiving. When caregiving responsibilities exceed ten hours per week, caregivers tend to allocate fewer hours to the labor market compared to non-caregivers (Lilly et al., 2007). Moreover, intense caregiving is predominantly undertaken by working-age women, who are often less likely to be fully employed and typically earn lower wages (Bauer and Sousa-Poza, 2015).

The consequences of informal caregiving, such as reduced employment rates and shortened work hours, lead to significant productivity losses for both employed caregivers and their employers. In Spain, caregivers of individuals with Alzheimer's disease experienced an average reduction of seven hours per week, equivalent to one lost workday in the last month or nearly two partial workdays (Darbà and Kaskens, 2015). The quantified productivity loss varied based on the level of impairment of the individuals with dementia (Gustavsson et al., 2011; Darbà and Kaskens, 2015; Michalowsky et al., 2016, 2018). For instance, a German study found higher monetary productivity losses in patients with mild and moderate cognitive impairment (Michalowsky et al., 2016, 2018). Additionally, in this German sample, caregivers' average weekly working hours decreased from 34 to 30 as the patients' cognitive impairments progressed. In Spain, the cost of productivity loss was estimated at 378 euros per month, although this figure included expenses related to the institutional setting, making direct comparisons challenging (Farré et al., 2018). A comparative study across Spain, Sweden, the United Kingdom, and the US, focusing on caregivers of community-dwelling Alzheimer's patients, revealed a substantial linear increase in productivity loss from mild to severe dementia, particularly evident in the Swedish sample (Gustavsson et al., 2011).

2.3 Institutional context in Mexico

Mexico provides a unique setting for our investigation into the interaction between health shocks and gender inequality, driven by several contextual factors.

First, Mexico's societal landscape is characterized by gender-based division of labor within households. Women traditionally bear a disproportionate burden of caregiving responsibilities (DiGirolamo and Salgado de Snyder, 2008), while men usually contribute a larger share of household income. These prevailing gender norms shape our inquiry, as they underlie gender inequality in income distribution and may contribute to disparities in the impact of partner's health shocks. Earlier studies noted that caregiving for the elderly often falls on female members of the household disproportionately.

Second, Mexico grapples with significant labor market informality (Aguilar-Gomez et al., 2019), which can be linked to a higher reliance on informal caregiving methods, such as within-household care. This informality is distinct from the United States, where formal healthcare is more prevalent. Understanding the dynamics of informality in Mexico is critical, as it might amplify the gender disparities in the consequences of spousal health shocks.

Moreover, Mexico exhibits pronounced occupational gender segregation, with women often concentrated in informal employment and sectors such as hospitality, which entail greater interpersonal interactions. This segmentation is relevant to our analysis as it can influence labor supply outcomes and the extent of reliance on household caregiving (World-Bank, 2019).

Additionally, the COVID pandemic has left an indelible mark on Mexico's economic landscape, intensifying the demand for caregiving and childcare due to lockdown measures. Globally, these responsibilities have disproportionately fallen on women, impacting their labor force participation and earnings (Alon et al., 2020). Our investigation seeks to disentangle the intricate interplay between spousal health shocks, traditional gender roles, healthcare disparities, and the informal labor market, all of which contribute to gender-based inequalities in the repercussions of health shocks.

3 Stylized Model

This study aims to explore household penalty, defined as the impact of a family member's infectious health shock on the health conditions and labor supply decisions of other household members. Our primary focus is the gender difference in household penalty.

We develop a one-period model of household health production based on Grossman (2000). Health depends on previous health stock and investment in this period. The former is taken as given, and the latter is a chosen level of optimization. Investment in health comes from care provided by an individual's family member. Health is a pure investment good and only indirectly affects utility through labor supply.

Computing household penalty relies on comparing the optimal decisions in infection-free households to the scenario where an infectious disease is present. We start by deriving optimal decisions for households without any infection. This serves as a benchmark. We then introduce an infectious disease into the household. In this new scenario, we re-evaluate the optimal responses of the household members to the presence of the infection. The *household* penalty is then quantified as the difference in outcomes between the infection-free benchmark and the scenario where the household is impacted by the infectious disease.

Model setup

Assume that a representative household includes a male and a female member, and they jointly consume an aggregated consumption good x_c.[1] They jointly maximize their household utility:

$$U = U[x_c]$$

The budget constraint is a function of wage, caregiving time, and health promotion income:

$$x_c \leq w_m(T - C_m) + w_f(T - C_f) + P(H_m - \bar{H}_0) + P(H_f - \bar{H}_0)$$

where w_m and w_f are wage rate for male and female. T is fixed total time endowments. C_m and C_f are time allocated to caregiving by male and female member, $1 \leq C_m \leq T$ and $1 \leq C_f \leq T$. Both of them get extra earnings if they have better health conditions, and the return rate is P. We have $w_m, w_f > 0$ and $P > 0$. The health production function is as follows:

$$H_m = \bar{H}_0 + \alpha ln C_f$$

$$H_f = \bar{H}_0 + \beta ln C_m$$

where $0 < \alpha < 1$, $0 < \beta < 1$. $\bar{H}_0$ is the same initial health stock for male and female. We assume care promotes health, and the marginal return of care is diminishing.

The household's problem is to maximizing household utility through their choices of C_m and C_f. This is equivalent to maximizing their household income. The first-order conditions are:

$$\left.\frac{\partial x_c}{\partial C_m}\right|_{C_m^*} = -w_m + \frac{\beta}{C_m^*}P = 0$$

$$\left.\frac{\partial x_c}{\partial C_f}\right|_{C_f^*} = -w_f + \frac{\alpha}{C_f^*}P = 0$$

The optimal level of each member's care and health is as follows:

[1]We develop a family bargaining model in Appendix section S1.1.

$$\frac{C_f^*}{C_m^*} = \frac{\alpha}{\beta} \cdot \frac{w_m}{w_f}$$

$$H_m^* - H_f^* = (\alpha ln\alpha - \beta ln\beta) + (\alpha - \beta) \cdot lnP - \alpha lnw_f + \beta lnw_m$$

Assumption 1: Gender wage gap

Assume that wage rate is different for male and female, and their production rate of providing care is the same. Namely we have $w_m > w_f$, 0< $\beta = \alpha$ <1. It yields:

$$\frac{C_f^*}{C_m^*} = \frac{w_m}{w_f} > 1$$

$$H_m^* - H_f^* = \alpha(lnw_m - lnw_f) > 0$$

Due to the gender wage gap, female member provides more care than male member, and female's health is worse than male's.

Assumption 2: Intra-household labor division

Now we close the wage gap and assume women have absolute advantage of providing care. We have $w_m = w_f$, 0< $\beta < \alpha$ <1, and the optimal level of care and health is:

$$\frac{C_f^*}{C_m^*} = \frac{\alpha}{\beta} > 1$$

$$H_m^* - H_f^* = (\alpha ln\alpha - \beta ln\beta) + (\alpha - \beta)(lnP - lnw)$$

Female member ends up providing more care than her partner due to the intra-household labor division. The sign of health condition gap is ambiguous. Female member has worse health condition than male if $lnw - lnP < (\alpha ln\alpha - \beta ln\beta)/(\alpha - \beta)$.

With both Assumption 1 and 2, together with the intra-household labor division, the gender income gap further makes female allocate more time to providing care to her partner, and the gender health gap is larger than before.

$$\frac{C_f^*}{C_m^*} = \frac{\alpha}{\beta} \cdot \frac{w_m}{w_f} > \frac{\alpha}{\beta}$$

$$H_m^* - H_f^* = (\alpha ln\alpha - \beta ln\beta) + (\alpha - \beta) \cdot lnP - \alpha lnw_f + \beta lnw_m$$

$$> (\alpha ln\alpha - \beta ln\beta) + (\alpha - \beta)(lnP - lnw)$$

Assumption 3: Infectious disease

So far, the only cost of providing care is the forgone income. Now consider a shock of infectious disease occurs in one family member. Providing care to the infected patient has a negative impact on the caregiver's health. Assuming the male is infected at the beginning, the new health production function is expressed as:

$$H_m = \bar{H}_0 + \alpha lnC_f - S$$

$$H_f = \bar{H}_0 + \beta lnC_m - \gamma lnC_f$$

where S is an exogenous shock due to the initial infection of an infectious disease. γ is the cost of providing care because the disease is infectious.[2] We assume $0 < \gamma < \alpha < 1$, otherwise no care will be provided to the infected member.[3] C_m^{S*} and C_m^* denote optimal care provided by male member with and without his getting infectious disease shock at the beginning, C_f^{M*} and C_f^* are optimal care provided by female member with and without her family member getting infectious disease shock at the beginning. The new optimal bundle of care and health is:

$$\frac{C_f^{M*}}{C_m^{S*}} = \frac{\alpha - \gamma}{\beta} \cdot \frac{w_m}{w_f}$$

$$H_m^{S*} - H_m^* = -S + \alpha lnC_f^{M*} - \alpha lnC_f^* = -S + \alpha[ln(\alpha - \gamma) - ln\alpha] < -S$$

$$H_f^{M*} - H_f^* = -\gamma lnC_f^{M*} = -\gamma ln\frac{(\alpha - \gamma)P}{w_f} < 0$$

This suggests that the male member's early infection makes the female provide less care than that in the infection-free condition due to the extra cost of caregiving. The second equation

[2]If the initial health shock is a non-infectious disease, we assume the health production function is the same as that in disease-free case except a shock term S. In other words, providing care to family member with non-infectious disease is not costly and care provider's health condition is not affected by care providing time. There is potential depression or psychological effect of family member's infection of non-infectious disease or infectious disease, which could be represented as an additional level change in care provider's health: $H_f = \bar{H}_0 + \beta lnC_m - \gamma lnC_f - D$, where D is the depression effect and is not affected by C_f, γ equals one if member's initial health shock is infectious and zero otherwise. The existence of depression effect or not does not change the optimization over C_m and C_f.

[3]We explore condition with $0 < \alpha \le \gamma < 1$ in Appendix section S1.2.

shows the negative effect on the male's health because the infection is amplified to some degree due to the female's lower care provision, shown as $H_m^{S*} - H_m^* > -S$. Due to her partner's infection, female member ends up with a worse health condition, shown as $H_f^{M*} - H_f^* < 0$. This means infection makes the health condition of other member in the same household worse than that in the infection-free condition, namely a household penalty.

If the initial health shock takes place on the female member, the household penalty on male is:

$$H_m^{M*} - H_m^* = -\gamma ln C_f^{M*} = -\gamma ln \frac{(\beta - \gamma)P}{w_m}$$

Back to our question of interest, we are comparing the health condition of male and female under the condition of other member's infection. Under Assumption 1, household penalty is larger for female due to gender wage gap:

$$\begin{aligned} H_m^{M*} - H_m^* &= -\gamma ln \frac{(\beta - \gamma)P}{w_m} \\ > H_f^{M*} - H_f^* = -\gamma ln \frac{(\alpha - \gamma)P}{w_f} &= -\gamma ln \frac{(\beta - \gamma)P}{w_f} \end{aligned}$$

Under Assumption 2, household penalty is larger for female due to absolute advantage of caregiving. In the following sections, we empirically quantify the household penalty for male and for female in the context of infectious disease.

$$\begin{aligned} H_m^{M*} - H_m^* &= -\gamma ln \frac{(\beta - \gamma)P}{w_m} \\ > H_f^{M*} - H_f^* = -\gamma ln \frac{(\alpha - \gamma)P}{w_f} &= -\gamma ln \frac{(\alpha - \gamma)P}{w_m} \end{aligned}$$

4 Data

4.1 Survey data

The Mexican Labor Survey (ENOE) is representative of the Mexican population and adopts a rotating panel design, conducting sequential household interviews spanning five quarters. An instrumental aspect of ENOE is its inclusivity, encompassing both the formal and informal labor sectors. This inclusivity is critical given the substantial prevalence of informal employment, a salient factor for our study, as our hypothesized mechanism may exert pronounced effects in the absence of formal health protection or insurance. The survey contains rich information regarding the household, its living conditions, its composition, all household members' socioeconomic characteristics, family relationships between them, and labor outcomes for members who are at least 15 years old.

Our dataset spans the years 2005 to 2019. To mimic the setting of the model and the insurance data, our primary sample concentrates on the households where both the head of the household and their partner are in the labor force. Moreover, we focus on the households where the couples are either in a formal or informal romantic relationship. We refine this by restricting our sample to individuals above age 18, yielding a dataset of approximately 23,000 households per quarter.

This dataset improves our study through multiple channels. Firstly, it equips us to evaluate the impact of health shocks on labor market outcomes. To identify a health shock, we consider an individual as "sick" if they were not working the previous week and attributed their absence to "illness, being excused, or personal affairs"[4]. To make sure that we consider new shocks rather than chronic illnesses, we only keep households which were healthy in the first quarter of the interviews. Contrary to the insurance claims data, in the survey data we cannot distinguish between who got sick first. We only observe whether a person was sick in the last week. Regarding the labor supply outcomes, we analyze how many hours the respondent worked last week.

Furthermore, this dataset provides occupation and family relationship details, enabling us to discern factors like the household composition, including the presence of potential alternative caregivers. These variables help to explore potential mechanisms underlying the results. Additionally, we have access to individual income data, enabling us to evaluate traditional models of labor supply within the family context and test model's predictions regarding income composition of the household.

[4]This does not include vacation, which are coded differently

4.2 Insurance claims

Second, we draw on US claims data, which provides detailed information on diagnosis codes, claim timing, and household members. This enables us to complement the analysis from Mexico by closely tracing how health conditions spread within households. In particular, we use de-identified health insurance claims processed by Office Ally and accessed through the COVID Research Database. The Office Ally is a clearinghouse that serves professional and institutional providers and processes their claims to both commercial and public payers (including Medicaid and Medicare). It includes up-to-date claims and remittance data in all states across the US since 2016. We use diagnosis code to filter claims of infectious disease based on the CDC's ICD-10-CM code.[5]

The most common type is respiratory infectious disease, followed by digestive, skin and soft tissue-related, bloodborne, and sexually transmitted infections. Detailed infectious disease list is reported in Table S22. COVID-19 is coded as Z20.828 in ICDâ10âCM and not included in our analysis. We also use complementary survey data to identify the top 10 most common health issues reported by women and men (table S23) and the top 10 health issues for which a medical professional was consulted (table S24). The gender differences in both are small. We are also able to observe patients' gender, year of birth, diagnosis date and healthcare providers' locations. Locations of patients are not available. The latest infection date in this study is August 1, 2020.

Besides, we use consumer data from AnalyticsIQ which creates a PeopleCore database using a blend of publicly available data and psychology-based algorithm. The database provides predicted characteristics of over 240 million individuals across 120 million households. Input data includes card purchase and online transactions that only covers publicly available information or opt-in data from consumers, vendors, and employees. The predicted data fields cover people's demographics, finance, credit, housing, jobs, lifestyle, behaviors, etc[6]. Household identifiers are predicted based on common budget. We obtain household-level variables from this database, including household identifiers, the number of children and adults, household income, and home value.

We use de-identified tokens to match individuals in these two datasets, and it gives us 1,125,908 individuals from 1,032,091 households in total. These individuals went to see a

[5]Available here: https://www.cdc.gov/nchs/icd/icd10cm.htm.
[6]Details on AnalyticsIQ could be found here: https://analytics-iq.com/

doctor at least once over the sample period, whether for infectious or non-infectious diseases. Given the construction process, patients in this study tend to come from middle and higher-income groups with health insurance. Among all the merged patients, we limited our focus to those whose households include at least one male and one female member. This practice excludes individuals who live alone or with same-sex family members from our sample. Altogether, there are 80,478 individuals from 38,105 households in 50 states and 645 three-digit zipcode areas in our sample.

5 Labor Penalty using Mexican Labor Survey

We begin by documenting the negative impact of a partner's illness on both health and labor outcomes in the context of Mexico. We find women exhibit a more pronounced deterioration in health and reduction in labor supply compared to men in response to their partners' illness. Moreover, our findings align with the model's predictions, showing that as the share of female income in the household increases, the female penalty diminishes while the male penalty amplifies. Additionally, in section S2 of the appendix, we analyze parents' labor supply responses to their children's sickness. We find that mothers' labor supply decreases significantly more (by 2.75 hours) compared to fathers' labor supply, following a child's health issues.

5.1 Empirical strategy

Our empirical strategy leverages partner's health shocks in the panel data to investigate the work-related outcomes within the working couples. We focus on two critical dependent variables: (1) whether an individual refrained from work in the previous week due to sickness or personal affairs (a binary outcome), and (2) the total number of working hours logged during the last week [7]. Our aim is to understand the interplay between gender, partner's sickness, and these outcomes.

$$\begin{aligned} Y_{it} = \beta_0 + \beta_1 \text{Sick Partner}_{it} + \beta_2 Female_i + \beta_3 Female_i \times \text{Sick Partner}_{it} \\ + \gamma_i + \delta_t + \varepsilon_{it} \end{aligned} \tag{1}$$

[7] In case of not working last week, this variable has value of 0

The estimated parameter β_1 captures the effect of partner's sickness on work outcomes for men. It quantifies the change in expected work outcomes of males when their partners are sick compared to when they are not. The parameter β_2 measures the gender effect on work outcomes, independently of partner sickness.

Our central hypothesis pertains to β_3, the interaction term $Female_i \times Member_i$. This parameter quantifies the additional effect of partner's sickness on work outcomes for women compared to the effect for men. Importantly, it tests whether there is a gender asymmetry in the household penalty in labor outcomes. Additionally, we include time fixed effects and, depending on the specification, municipality or individual fixed effects.

Our empirical strategy compares changes in outcomes over time for individuals whose partners become sick to those whose partners remain healthy. Naturally, households where an individual becomes sick might differ from those where all members stay healthy. As shown in the balance table S7, during the initial interview period when all households are healthy, there are small but statistically significant differences between the two groups. To address this, we reweight the sample to match observable characteristics between the groups and obtain qualitatively similar results (see table S2). Importantly however, for our strategy to remain valid, it is not necessary for sick and healthy households to be identical at baseline when it comes to levels. We only require that, in the absence of the health shock, the outcomes for sick households would evolve in parallel to those of healthy households.

To validate our identifying assumptions, we perform an event study analysis leveraging the panel structure of the data, where each household is observed over five quarters with an interview in each period. This approach allows us to assess potential pretrends by comparing households in which a partner becomes sick (in periods 3, 4, or 5) to households where no illness occurs. Specifically, we estimate the following specification separately for men and women:

$$Y_{it} = \sum_{k=-4}^{2} \beta_k \{\text{Distance from event=k}\}_{i,t}^{k} + \gamma_i + \delta_t + \varepsilon_{it}$$

where Y_{it} is the outcome of interest for individual i in period t, $\text{EventTime}_{i,t}^{k}$ are indicators for time relative to illness onset ($k = 0$ indicates the onset), γ_i denotes individual fixed effects, δ_t represents time fixed effects, and ε_{it} is the error term.

To obtain percentage interpretation we estimate this model using Poisson regression, but also using ordinary least squares (OLS) on levels. The coefficients $\beta_{-4}, \beta_{-3}, \beta_{-2}$ test for potential pretrends, with the absence of significant differences supporting the parallel trends assumption. The coefficient β_0 measures the immediate impact of the illness. Estimating this model separately by gender further allows us to investigate whether the effects of a partner's illness and any pretrends differ systematically between men and women.

Next, we extend our analysis to understand the potential mechanisms driving the gender assymetry in the household penalty. We incorporate heterogeneity factors (represented by the HT term), which encompass variables related to the presence of an elderly woman in the household, and the share of female income. The extended regression equation is as follows:

$$\begin{aligned}
Y_{it} &= \beta_0 + \beta_1 \text{Sick Partner}_{it} + \beta_2 Female_i + \beta_3 HT_i \\
&+ \beta_3 Female_i \times \text{Sick Partner}_{it} + \beta_4 Female_i \times HT_i + \beta_5 \text{Sick Partner}_{it} \times HT_i \\
&+ \beta_6 Female_i \times \text{Sick Partner}_{it} HT_i \\
&+ \gamma_i + \delta_t + \varepsilon_{it}
\end{aligned} \tag{2}$$

These factors are relevant for working couples' dynamics and household composition. Our hypotheses are as follows:

The presence of an elderly female (age>60) in the household is hypothesized to provide additional potential caregiving support. We expect households with an elderly female member to be better insulated from the negative consequences of partner's sickness, potentially reducing the impact on women's work outcomes.Given the potential pre-shock differences between households with and without elderly female presence (see balance table S3), we re-estimate the model using a reweighted sample, ensuring that households with and without elderly presence are, on average, similar on observables.

Regarding the share of female income, we test our model implication that the household specialization determines the optimal care-giving level. A larger share of female income should lead to men being more likely to reduce working hours or take sick leave when their female partners are sick, reflecting comparative advantage within the household. Conversely, women, as primary income earners, are less likely to reduce their working hours or take sick leave when their male partners are ill, as it would disrupt the primary source of household income.

New parameters of interest shed additional light on the mechanisms of household penalty for women. β_4 quantifies how gender ($Female_i$) moderates the influence of the heterogeneity factor (HT_i) on work outcomes. It helps to discern whether the effect of variables like the presence of an elderly woman, or the share of female income differs between men and women within working couples.

β_5 represents the interaction between partner's sickness ($Member_i$) and household dynamics (HT_i). It highlights how specific household characteristics modify the impact of partner's sickness on the work outcomes of men. β_6 is the critical parameter of interest as it examines whether the heterogeneity term drives the gender asymmetry in the household penalty. Namely, it estimates how our heterogeneity terms mediate the additional impact of male sickness on female outcomes. It is important to note that the variation in the heterogeneity factors is not exogenous. Therefore, this heterogeneity analysis is descriptive rather than causal. We also supplement this analysis with event-studies where sample is split based on the value of the heterogeneity factor

5.2 Main results

The results from estimating Equation (S.4), as presented in Table 1, provide substantial evidence of greater health and labor-related penalties for women. We focus on the preferred specification with individual fixed effects. Specifically, in Panel A, we examine the likelihood of being 'Sick'. We observe a notable gender disparity: when the female partner in a working couple falls sick, there is a 8.1 percentage point increase in the probability of her male partner also being sick. In contrast, if the male partner is sick, the likelihood of the female partner also being sick escalates by 13.2 percentage points. These differences are not only statistically significant but also economically meaningful, especially when considered in the context of the average probability of missing work in a given week, which stands at 1.2%. This effect is further illustrated in the event study graph shown in Figure S3. Although minor deviations from the trend are observed prior to the event, they are negligible compared to the magnitude of the main effect.

The "Hours Worked Last Week" outcome exhibits a similar gender asymmetry as shown in Panel B. In the main specification (Column (2)), we find that when the female partner is sick, the male partner works 4.2 hours less on average, representing a 10% decline from the average weekly working hours for men, which is 44.9 hours. These negative spillovers are even more pronounced for women, providing further evidence of the household penalty. The interaction term indicates that when the male partner is sick, the female partner reduces her labor supply by 4.7 hours. This reduction is particularly significant given the average weekly working hours for women is 34 hours, translating to a 15% decline. The difference in hours worked between male and female reactions is significant at the 5% level.

The event study results corroborate the findings from the difference-in-differences analysis, showing no significant deviations from parallel trends prior to the event. Figure 1 presents the results estimated using a Poisson model, while Figure S2 illustrates the estimates in levels.

While the overall differences between genders are modest, this can largely be attributed to the fact that women, on average, work fewer hours than men, leaving less scope for a substantial decline. However, when restricting the analysis to individuals working full-time (40 hours or more) during the pre-event periods, the disparity becomes more pronounced. Female working hours decrease by 6.35 hours compared to a 4.34-hour reduction for males (Figure S4).

This setting inherently combines two effects: (i) a reduction in working hours due to the individual becoming sick, potentially as a result of contagion from their partner, and (ii) a reduction in hours driven by increased caregiving responsibilities. To disentangle these effects and isolate the caregiving impact, we analyze a subsample of individuals who did not become sick when their partner fell ill. In this context, we still observe a decline in hours worked, albeit smaller, averaging approximately one hour (Figure S5) or slightly more than 2% (Figure S6). Notably, this reduction does not exhibit gender asymmetry, suggesting that the additional female penalty largely arises from women's higher likelihood of becoming sick when their partner is ill.

5.3 Heterogeneity results

In this section, we perform the heterogeneity analysis to explore the mechanisms driving the asymmetric gender penalty. We examine how it varies based on the presence of elderly women in the household, income shares within the couple and whether the individual works in formal or informal sector.

First, there is a limited evidence which suggests that the presence of an elderly female attenuates the female penalty for the working woman in the household. Panel A of Table 3 presents the results for both the full and reweighted sample are presented. The triple interaction term captures the change in female penalty for being sick when another elderly female is present in the household. The full sample results tell us that the additional female penalty is 6 percentage points smaller when an elderly female is present, which is consistent with the insulating role of "grandmas". Nonetheless, this effect disappears in the matched sample. In Panel B, the coefficient on the triple interaction term is positive, indicating that the reduction in female working hours is smaller when an elderly female is present. While the magnitudes suggest that female penalty does not exist when elderly female is present, coefficients are not statistically significant.

We complement this analysis with an event study using Poisson regression to explore the differential evolution of outcomes for the sample of individuals living with an elderly female. When a partner becomes sick, women reduce their working hours by approximately 16.4%, while men reduce theirs by 15.9%. This represents a much smaller gender gap compared to the overall sample and is not statistically significant (Figure S9). Interestingly, the presence of an elderly male in the household does not appear to confer a similar protective effect for women. In households with an older male, women reduce their hours by 18%, whereas men reduce theirs by only 10%, resulting in a significant gender disparity (Figure S10). These findings are consistent with a narrative where caregiving responsibilities disproportionately fall on women.

Our analysis, emphasizing the female share of household income, strongly supports the theory of household specialization as a primary factor in the additional female penalty. As detailed in Panel A of Table 3, we focus on the interaction term representing the percentage of income contributed by women in a couple. The coefficient's magnitude indicates a significant shift: in households where the woman is the sole earner, the additional female penalty in terms of sickness is completely erased. On the other hand, males are more likely to fall ill when a woman is sick when the female share of income is high. Hence, the household penalty is inverted in households where women are primary earners. This is a notable result which differentiate this setting from motherhood penalty where the share of female income does not matter (ex: Almond et al. (2023)).

Panel B presents evidence supporting household specialization in terms of its influence on working hours. When we account for individual fixed effects , we observe a positive and significant triple interaction effect. If a woman is the sole earner in the household, it completely offset the reduction in work hours that women typically experience when their partners have health issues. However, this result is not consistent across all specifications. Moreover, we do not find the effect of income share on male reaction to partner's illness. Overall, these findings partly align with economic theories about how households allocate labor based on who can do the job most efficiently, and they suggest that households adapt their labor strategies based on their comparative advantage. In conclusion, we find that the the presence of an elderly female, and the higher female income share insulate women from the penalty.

Figures S7 and S8 illustrate how the response to a partner's illness differs between workers in the formal and informal sectors. The formal sector's rigid work arrangements limit adjustments to working hours, amplifying societal expectations that women reduce work for caregiving. In contrast, the informal sector's flexibility allows both men and women to adjust their hours more easily, reducing the role of gender norms.

Dividing the sample by sector, we find a larger gender gap in the formal sector. Women in the formal sector reduce their hours by 8% (3.15 hours), while men reduce theirs by 5% (2.36 hours), reflecting smaller overall declines due to greater rigidity. In the informal sector, women reduce hours by 9% (2.6 hours) and men by 8% (3.4 hours), with narrower gender differences percentage-wise.

The analysis of the Mexican survey data aligns closely with the model's predictions: having a sick partner negatively affects health, and this impact extends to labor supply. Furthermore, consistent with the model, the gender income gap significantly contributes to the asymmetric burden, placing women in a considerably disadvantaged position. While the ENOE data clearly demonstrates the presence of the household penalty, it does not allow us to determine whether this effect, as the model suggests, operates through the transmission of infectious diseases within the household. To explore this channel more precisely, we turn to insurance claims data.

6 Health Penalty using US Insurance Claims

We examine the concept of "household penalty" in the context of health and infectious disease transmission by utilizing insurance claims data from the United States.

6.1 Empirical strategy

To quantify the household penalty and its difference by gender in the insurance claims data, we use a standard difference-in-difference design:

$$\begin{aligned} Y_{ih} = \beta_0 + \beta_1 Member_i + \beta_2 Female_i + \beta_3 Female_i \times Member_i \\ + \beta_4 Age_i + \beta_5 Age_i^2 + \beta_6 HHSize_h + ZIP_h + \varepsilon_{ih} \end{aligned} \tag{3}$$

where the sample includes 80,478 patients with available claims, gender, and age. Their household information, including household identifier, income, and house sizes, is observed at each diagnosis. We exclude one-individual households from the sample. $Member_i$ is an indicator that equals one if individual i has an infected family member in the household that has been diagnosed with an infectious disease, captured by diagnosis code. $Member_i$ is interpreted as the treatment of household contagion exposure, whether i is exposed to household infection or is only exposed to infection outside households. Consider two types of households in our sample: infected households and infection-free households. All individuals in infection-free households, as well as the very first infected members in infected households, are assigned $Member_i$ equal to zero because the first patient is not exposed to other members' infection. They are considered the control group without household contagion. All individuals except the first patient in infected households are assigned $Member_i$ as one, and they comprise the treated group.

Outcome Y_i is an indicator that equals one if individual i is infected by the same type of infectious disease as his/her family member's within one year, and infection timing is captured by diagnosis date. We use a one-year cutoff to balance the time period because a shorter duration may fail to capture the recovery period and transmission within household members, while a longer duration suggests that others' infections are less likely due to household transmission.[8] Results using alternative cutoff periods are reported in Section 6.2.2. Regarding the same type of infectious disease, they are coded based on the same letter and first two digits using the ICD-10-CM code. Within the treated group, those infected one year later than their members or those never infected in an infected household have Y_i equal to zero. Within the control group,

[8] Specifically, a long post-infection period ensures that all infected members are included in the outcome variable, accounting for infections that may not occur immediately after a family member's infection and may have a long incubation period. Additionally, some diseases exhibit seasonality, and a short post-infection period might coincide with seasonal cycles of the disease, potentially biasing our estimates.

all patient zeros are assigned Y_i as one and those from infection-free households have Y_i as zero. The control group is used to construct the baseline infection rate without household contagion exposure and its gender and geographic differences.

$Female_i$ is an indicator of being a female. Coefficient β_2 is the gender difference of infection rate due to biological vulnerability, protective behaviors or occupation difference and has nothing to do with family transmission. Coefficient β_1 is the extra probability of getting infected due to an infected family member, the so-called household penalty. It captures the difference in the occurrence of infectious disease between the control and the treated group, namely between those facing only outside contagion and those facing both household contagion and outside contagion. The interaction term $Female_i \times Member_i$ captures the additional household penalty for women. The household penalty is estimated as β_1 for males and $\beta_1 + \beta_3$ for women.

Additionally, we add quadratic age terms, household size in a linear term or fixed effects, household income and home value on the right-hand side to control other characteristics that may affect the probability of infection. To control for unobservable geographic differences, we add three-digit zipcode fixed effects using the healthcare provider's location.

Our identifying assumption is that gender difference is the same for the treated and the control group if the household transmission channel is closed. In other words, the observed gender difference between these two groups solely comes from household contagion. Family member's infection may increase awareness, and people may switch jobs to avoid inter-personal interactions or adopt more protective measures outside households, but this response is less likely to differ by gender.

Since our data on infections are derived from insurance claims, they reflect instances where individuals have sought medical attention. It is important to acknowledge that differential health-seeking behavior between men and women might partially influence our results. However, we provide suggestive evidence that such behavior accounts for only a small fraction of the observed gender disparity. Specifically, in Appendix section S3, we demonstrate that the magnitudes of our findings are inconsistent with the gender gap arising solely from differences in the propensity to visit the doctor, and instead reflect underlying differences in infection patterns.

Besides, the insurance dataset may not be able to capture all the visits for infectious disease. We assume the latent probability is the same for the treated and the control group: namely the capability of capturing infectious disease visits in the claim data is the same for these two groups.

6.2 Main results

6.2.1 Summary statistics

Table S5 shows some summary statistics of the outcome and control variables. Individuals in our sample tend to be middle-aged, with an average of 59 years old. We don't observe any child in our sample, so child's infection is excluded when we code the first infection in household or outcome variable, and caregiving to child is not the channel we explore. There are slightly more women than men, 50.6% of the total. The likelihood of a family member's earlier infection is 11.85%. Namely 11.85% of individuals in our sample face household contagion and are considered the treated group, while the rest 88% only face contagion outside households. The control group includes all individuals in infection-free households and all the first infected member in infection households. The mean of the outcome variable is 9.7%. The outcome variable captures the occurrence of infectious disease anytime within our sample period for the control individuals, and the occurrence of infectious disease within one year of a member's infection for the treated individuals. Thus, 9.7% is the weighted average of these two infection rates and is slightly smaller than the infection rate due to the restricted definition for the treated group. However, it is similar to the simple infection rate since the control group is six times larger than the treated group.

In Table 4, the first three rows show the average outcome variable for all individuals, men and women regardless of member's infection. Taking all individuals together, women are more likely to be infected than men, 10% vs. 8.5%, though the infection rates are not significantly different. Under the condition of no household contagion, the likelihood of being infected slightly goes down, and its gender difference is similar to that of the unconditional mean. In contrast, the middle three rows show the probability of infection increases from 9.3% to 12.9% after a member's infection. Separating member's gender, women are much more likely to be infected when the early infection takes place in male member than that in female member, 14.4% vs. 12.9%. For men, their infection rates remain similar with and without a male member's infection, 8.6% and 8.5%, but increase to 12.1% after a female member's infection.

The descriptive evidence shows there are significant male-to-female, female-to-female, and female-to-male transmission, but no obvious male-to-male transmission.

6.2.2 Household penalty and gender difference

Regression results from estimating equation (S.3) are reported in Table 5. Positive coefficients on $Member$ capture the impact of household contagion on other member's infection within one year. This confirms the higher likelihood of infection after a family member is infected. The coefficient remains similar after adding linear household size, household size fixed effects (9 dummies), household income, and house value on the right-hand side. The infection rate of individuals with infected member is higher by 1.2 percentage points compared with those without member's earlier infection. This household penalty is equivalent to 12.4% relative to the mean infection rate. Besides, the positive estimate of coefficients on $Female$ suggests women are more likely to be infected regardless of member's infection. This may be due to gender occupational difference and its associated exposure to contagion (Lewandowski et al., 2020) or different biological vulnerability. The coefficient of interest on the interaction term $Member \times Female$ is significant and positive. This indicates the household penalty is even higher for female members after controlling for the baseline gender infection difference. Due to household contagion, women are 2.2 percentage points more likely to be infected, 22.7% relative to the mean. The results are robust with and without household characteristics.[9]

To disentangle the effect of household size, in Table S8, we replace household size with the number of children and the number of adults in the household. Coefficients on $Member$, $Female$, $Member \times Female$, and Age are quite similar to those in Table 5. Estimates on the new covariates show the probability of infection decreases with the number of adults in the household and is not significantly different for those living with a large or small number of children. This suggests that close contact among adult members contributes to intra-household disease transmission.

Apart from infection within one year, we also use infection 1-2, 2-3, and 3-4 years after a

[9] In terms of other covariates, the estimates on Age and Age^2 show a nonlinear relationship between infection and age and the least vulnerable age is estimated to be 52 years old. Compared with the least vulnerable group, the infection rate of people who are 62 and 72 years old is 0.53 and 2.1 percent higher. Regarding household characteristics, the estimate on household size is negative but imprecise. When household income and home value are added in the regression, R^2 slightly increases from 0.353 to 0.357 and lower home value is associated with more infection. The results are consistent with previous findings that older people and the low income group are more seriously attacked by infectious diseases (Belot et al., 2020; Wiemers et al., 2020).

member's infection to code the outcome variable. Results in Table S9 show imprecise estimates on $Member$ and $Member \times Female$ in all panels. There is a similar baseline gender infection difference as that in Table 5. The estimated household penalty is positive for 1-2 years' infection, but the estimate is not significantly different from zero and the magnitude is much smaller as the time gap increases. The positive estimates indicate the effect of a member's infection on others infection 1-4 years later, and it could result from the spillover of a member's effect within one year. Figure S11 displays the increased infection rate for the treated group for men and women, and the gender difference of increased infection. Similar to regression results, we only find household penalty within one year and no more penalty later. There is no gender gap in the increased infection rate after one year either. This suggests member's infection has no impact on other members' health outcomes after one year, neither does its gender difference.

6.2.3 Within household comparison

We conduct a within-household comparison of gender difference only using individuals in the treated group with household contagion. We further require there be at least a male and a female member left after the first patient is dropped. Instead of a double-difference design, we conduct a single-difference analysis, drop $Member$ and the interaction term, and add household fixed effects. Other controls are the same as those in equation S.3. Table S10 shows a higher infection rate faced by women by 2.7 percentage points. The result shows a similar disproportionate household infection faced by women.

In the main specification without household fixed effects, we assume member's infection event does not change the baseline gender infection difference after closing household transmission channel. We explore the across-household within-zipcode variation and compare the gender difference in the treated and the control group. In this section, the within-household comparison shows female's disproportionate infection in comparison with her own male family member. The estimated gender difference is slightly larger. Though this practice results in a smaller sample size, larger standard errors and a restricted household structure with three or more members, we favor the within-household comparison to absorb any unobservable characteristics of households.

6.3 Heterogeneity results

In this section, we examine heterogeneous effects across family members' gender and age. We also explore heterogeneities by race and job in Appendix Section S4.

6.3.1 Heterogeneous effects by member's gender

The observed household penalty comes from the impact of any family member's early infection, regardless of the member's gender. In this section, we separately use male and female member's infection to code $Member$ and re-estimate equation (S.3). When analyzing the impact of male member's infection, the control group is exactly the same as before, i.e., individuals in infection-free households and all the household primary infections regardless of gender. For the treated group, we exclude households where only female members are infected and retain those exposed to contagion from male members. This results in a different sample size for studying the impact of men's and women's earlier infections, and a different sample size compared to the pooled results in Table 5.

Results in Table S11 show imprecise estimates on $Member_m$, which suggests the impact of other male member's early infection is not significant on male members. Namely there is no significant evidence of male-to-male transmission of infectious disease within households. The positive and significant estimates on $Member_m \times Female$ show the secondary attack rate of male-to-female transmission is 2.8 percentage points or 28.7% relative to the mean. An early infection in a male member only affects female members in households. This is consistent with the conditional mean in Table 4 and is likely to result from female's caregiving to infected patients.

When using a female member's infection to code $Member$, results in Table S12 show the same secondary attack rates for men and women. After women's infection, they may still need to do domestic work and provide care to other members, and the exposure to infection faced by other members is not significantly different by gender. Taking all family members together, the observed extra household penalty faced by women is driven by the significant male-to-female and insignificant male-to-male transmission. Besides, the likelihood of female-to-male infection is estimated to be 1.5 percentage points, smaller than that of male-to-female infection.

We also use the same sample as that in Table 5 and separate the first patient's gender to code $Member$. Namely we use two variables to replace $Member$ and add two interaction terms.

Results in Table S13 show similar transmission patterns similar to those in Table S11 and S12. When the first infection in household is male, the increased infection rate for the other men and for women is 0 and 4.2 percentage points, respectively. Symmetrically, female's infection as the first patient in household increases men's and other women's infection rates by 3.3 percent.

6.3.2 Heterogeneous effects by member's age

Apart from member's gender, we also explore the heterogenous effects by the first patient's age. To do so, we separate individuals into three age groups and replace $Member$ with three dummies given the first patient's age. Results in Table S14 display the impacts of infected member with 40 and 60 years old as cutoffs.

Estimates on $Member_{old}$ and $Member_{old} \times Female$ show an infected member over 60 years old increases other male's and female's infection rate by 4.3 percentage points. There is no gender difference in the secondary attack rates when the first patient is old. Similar patterns are found when the first infected is in his/her 20s and 30s. Both men and women face a higher infection rate by 3.5 percentage points, slightly lower than that from an old member's early infection.

In contrast, women face unequally higher infection risk when the first patient is 40-60 years old. A potential explanation for this is that middle aged members are healthy in general. When they are infected, they need some care but not a lot. Not all the rest members are needed to care for them, and female member will take on this burden. In contrast, the old group may require intensive care provided by the rest of the members, and this results in higher exposure for both male and female. Another potential reason is that, the middle aged group contribute more to the household income. When they are infected, the loss of household income is high. The other members need to make up for the household income, and they don't want to lose one male worker with a higher income. So only female is providing care to this middle-aged patient.

In Table S15, we further separate the outcome individual's age with the same cutoffs, 40 and 60, to check subgroup transmission patterns separately. Focussing on middle-to-young and old-to-middle transmission, the burden of taking care of older patients is unequally shared by women, captured by the positive estimates on $Member_{middle} \times Female$ in Column (1) and $Member_{old} \times Female$ in Column (2). In contrast, the burden of taking care of younger patients is equally shared by men and women. The disproportionate household transmission faced by women in the pooled results is driven by women taking care of older patients. Apart from the intergenerational transmission,[10] the within-age transmission is only significant in the old group and old couples are likely to catch the disease together. Among the six groups, young and middle-aged women below their 40s face the highest risk of getting infected.

6.4 Mechanisms

In the insurance data, we find again patterns consistent with the asymmetric household penalty being driven by household specialization. Specifically, given the two assumptions in Section 3, we test the heterogeneous patterns of household penalty by different gender income gap and different home production intensity.

First, the female penalty is larger in the areas with a higher gender income gap. We use income by gender data from the American Community Survey 2016 and calculate the gender income gap for each zipcode area.[11] Figure S1 Panel A plots spatial distribution of gender wage gap. Then we separately estimate equation (S.3) for zipcode areas above and below the median and results are shown in Table S16. As is shown in Panel A, for individuals in areas with small gender income gap, the household penalty is still positive and large but estimates are imprecise. Moreover, the penalty is not significantly different by gender. The small and imprecise estimates on the interaction term suggest both men and women have a similar probability of infection with a family member's early infection. In contrast, the positive estimates on $Member$ and $Member \times Female$ in Panel B show similar patterns as those in Table 5. Men have a 1.2 percentage points higher infection rate and women have a 2.4 percentage points higher rate. The observed results taking all areas together are mainly driven by areas with large gender income gap.

Second, the female penalty is larger in the areas where women spend more time on home production. On the second assumption, we use data from the American Time Use Survey to construct the gender difference of home production intensity at the zipcode level.[12] Figure S1

[10] In our sample, 26.2% of individuals live in intergenerational households where the gap between min and max age is over 20 years old.

[11] The geographic unit in the American Community Survey is the census block group. We use male and female's average income for each census block group, and calculate the population-weighted average male and female income for each zipcode area. Then we define gender income gap as male's income minus female's over male's, and separate zipcodes by the median gap.

[12] We use data after 2014 with precise county information and calculate the weighted average male and female's time allocated to home production. Then we calculate gender time gap as female's time minus male's over male's, and separate zipcodes by the median gap.

Panel B plots spatial distribution of gender time use difference. Table S17 displays estimation results using areas with smaller or larger gender gap. While women always face higher household penalty than men, in areas with a higher gap in home production intensity, men are not affected by members' earlier infection. This indicates that men in areas where they face a relatively lower home production burden than women do not suffer from member's earlier infection. The gender gap in home production intensity drives the observed disproportionate household transmission and is likely to be the second mechanism.

6.5 Robustness

For robustness, we relax the requirement when constructing our sample. Instead of requiring at least a male and a female, we require two individuals in households regardless of their gender, and the new sample size is twice as large. In Table S20 Panel A, we find similar estimates on $Member$ and $Member \times Female$ to those in Table 5. Men and Women are 1.5 percent and 2.5 percent more likely to be infected due to household contagion exposure.

Moreover, we use a subsample with no baseline gender infection difference and re-do the estimation in case our results are driven by the higher probability of women seeing doctors than that of men. Results in Table S20 Panel B remain robust, and the household penalty for men and women are estimated to be 2.2 and 3.2 percent, 22.4% and 32.5% relative to the mean.

As a placebo test, we use the infection of cancer (a non-infectious disease) to replace the outcome and $Member$. We find no effect of member's early infection on another member's infection and no gender difference, as is shown in Table S21.

7 Conclusion

This paper studies the spillover effects of health shocks on family members' health conditions and labor outcomes. Our empirical analysis covers both developing and developed countries, revealing the generalizability of household spillovers and the disproportionately adverse consequences faced by female household members.

We propose a concept, household penalty, and construct a theoretical framework on infectious disease transmission within households. One mechanism of gender inequality comes from the unequal distribution of household labor, with women consistently shouldering a heavier burden of housework and childcare responsibilities. In the context of infectious diseases, women's caregiving roles become even more pronounced, with their contributions taken for granted as they provide care to infected family members. Furthermore, the disproportionate infection rates among Women may be a consequence of their additional housework burden due to other members' infections. With a smaller pool of household members available to perform housework, women are often left to shoulder a greater share of the remaining responsibilities, indirectly exacerbating their vulnerability.

Another potential mechanism lies in the income disparity between genders. If men within the household have higher incomes or serve as the primary breadwinners, the family may collectively decide to prioritize the protection of male members from infection or expedite their recovery if infected.

This paper bridges several fields of economics research, including the economics of household, health, labor, and gender economics. Our analysis underscores the importance of combining fields in future research. Besides, access to longer-term health records, especially those related to the COVID pandemic and seasonal influenza, will prove invaluable in expanding the sample size and deepening our understanding of the health crisis.

References

Aguilar-Gomez, S., E. Arceo-Gomez, and E. De la Cruz Toledo (2019). Inside the black box of child penalties: Unpaid work and household structure. *Available at SSRN 3497089*.

Almond, D., Y. Cheng, and C. Machado (2023). Large motherhood penalties in us administrative microdata. *Proceedings of the National Academy of Sciences 120*(29), e2209740120.

Alon, T., M. Doepke, J. Olmstead-Rumsey, and M. Tertilt (2020). The impact of covid-19 on gender equality. Technical report, National Bureau of economic research.

Andrew, A., S. Cattan, M. Costa Dias, C. Farquharson, L. Kraftman, S. Krutikova, A. Phimister, and A. Sevilla (2020, July). The gendered division of paid and domestic work under lockdown. *IZA Discussion Paper* (13500).

Archibong, B. and F. Annan (2017, May). Disease and gender gaps in human capital investment: Evidence from niger's 1986 meningitis epidemic. *American Economic Review 107*(5), 530–35.

Archibong, B. and F. Annan (2020). Schooling in sickness and in health: the effects of epidemic disease on gender inequality. *CDEP-CGEG Working Paper* (54).

Arrieta, G. R. and G. Li (2023). Caring to work or working to care: The intra-family dynamics of health shocks. *American Journal of Health Economics 9*(2), 175–204.

Barreca, A., K. Clay, O. Deschenes, M. Greenstone, and J. S. Shapiro (2016). Adapting to climate change: The remarkable decline in the us temperature-mortality relationship over the twentieth century. *Journal of Political Economy 124*(1), 105–159.

Bauer, J. M. and A. Sousa-Poza (2015). Impacts of informal caregiving on caregiver employment, health, and family. *Journal of population Ageing 8*, 113–145.

Belot, M., S. Choi, E. Tripodi, E. van den Broek-Altenburg, J. C. Jamison, and N. W. Papageorge (2020, June). Unequal consequences of covid-19 across age and income: Representative evidence from six countries. *IZA Discussion Paper* (13366).

Brito, E. and D. Contrera (2023). The caregiving penalty: Caring for sick parents and the gender pay gap. *Working Paper*.

Chan, J. F.-W., S. Yuan, K.-H. Kok, H. Kai-Wang To, Kelvin Chu, J. Yang, F. Xing, J. Liu, C. C.-Y. Yip, R. W.-S. Poon, H.-W. Tsoi, S. K.-F. Lo, K.-H. Chan, V. K.-M. Poon, W.-M. Chan, J. D. Ip, J.-P. Cai, V. C.-C. Cheng, H. Chen, and C. K.-M. Hui (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. *The Lancet 395*(10223), 514–523.

Chernoff, A. W. and C. Warman (2020, July). Covid-19 and implications for automation. Working Paper 27249, National Bureau of Economic Research.

Coile, C., M. Rossin-Slater, and A. Su (2022). The impact of paid family leave on families with health shocks. Technical report, National Bureau of Economic Research.

Curmei, M., A. Ilyas, O. Evans, and J. Steinhardt (2020). Estimating household transmission of sars-cov-2. *medRxiv*.

Darbà, J. and L. Kaskens (2015). Relationship between patient dependence and direct medical-, social-, indirect-, and informal-care costs in spain. *ClinicoEconomics and Outcomes Research*, 387–395.

Daysal, N. M., H. Ding, M. Rossin-Slater, and H. Schwandt (2021). Germs in the family: The long-term consequences of intra-household endemic respiratory disease spread. Technical report, National Bureau of Economic Research.

DiGirolamo, A. M. and N. Salgado de Snyder (2008). Women as primary caregivers in mexico: challenges to well-being. *salud pública de méxico 50*(6), 516–522.

Esteve, A., I. Permanyer, D. Boertien, and J. W. Vaupel (2020). National age and coresidence patterns shape covid-19 vulnerability. *Proceedings of the National Academy of Sciences 117*(28), 16118–16120.

Fadlon, I. and T. H. Nielsen (2019). Family health behaviors. *American Economic Review 109*(9), 3162–3191.

Fadlon, I. and T. H. Nielsen (2021). Family labor supply responses to severe health shocks: Evidence from danish administrative records. *American Economic Journal: Applied Economics 13*(3), 1–30.

Farre, L., Y. Fawaz, L. Gonzalez, and J. Graves (2020, July). How the covid-19 lockdown affected gender inequality in paid and unpaid work in spain. *IZA Discussion Paper* (13434).

Farré, M., B. Kostov, J. M. Haro, E. Cabrera, E. Risco, M. Alvira, S. Miguel, and A. Zabalegui (2018). Costs and burden associated with loss of labor productivity in informal caregivers of people with dementia. *Journal of Occupational and Environmental Medicine 60*(5), 449–456.

Galasso, V., V. Pons, P. Profeta, M. Becher, S. Brouard, and M. Foucault (2020, June). Gender differences in covid-19 related attitudes and behavior: Evidence from a panel survey in eight oecd countries. Working Paper 27359, National Bureau of Economic Research.

Gimenez-Nadal, J. I., J. Alberto Molina, and J. Velilla (2020, May). Should we cheer together? gender differences in instantaneous well-being during joint and solo activities. *IZA Discussion Paper* (13306).

Glynn, J. R., H. Bower, S. Johnson, C. Turay, D. Sesay, S. H. Mansaray, O. Kamara, A. J. Kamara, M. S. Bangura, and F. Checchi (2018). Variability in intrahousehold transmission of ebola virus, and estimation of the household secondary attack rate. *The Journal of Infectious Diseases 217*, 232–237.

Goldin, C. and A. Lleras-Muney (2018, June). Xx>xy?: The changing female advantage in life expectancy. Working Paper 24716, National Bureau of Economic Research.

Grossman, M. (2000). The human capital model. In A. J. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics* (1 ed.), Volume 1, Chapter 07, pp. 347–408. Elsevier.

Gustavsson, A., P. Brinck, N. Bergvall, K. Kolasa, A. Wimo, B. Winblad, and L. Jönsson (2011). Predictors of costs of care in alzheimerâs disease: a multinational sample of 1222 patients. *Alzheimer's & Dementia 7*(3), 318–327.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy 126*(S1), S197–S246.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings*, Volume 109, pp. 122–126. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Lee, S. (2020). Spousal labor supply, caregiving, and the value of disability insurance. *Caregiving, and the Value of Disability Insurance (April 30, 2020)*.

Leslie, E. and R. Wilson (2020). Sheltering in place and domestic violence: Evidence from calls for service during covid-19. *Journal of Public Economics forthcoming*.

Lewandowski, P., K. Lipowska, and I. Magda (2020, June). The gender dimension of occupational exposure to contagion in europe. *IZA Discussion Paper* (13336).

Li, W., B. Zhang, J. Lu, S. Liu, Z. Chang, C. Peng, X. Liu, P. Zhang, Y. Ling, K. Tao, and J. Chen (2020, 04). Characteristics of Household Transmission of COVID-19. *Clinical Infectious Diseases*. ciaa450.

Lilly, M. B., A. Laporte, and P. C. Coyte (2007). Labor market work and home care's unpaid caregivers: a systematic review of labor force participation rates, predictors of labor market withdrawal, and hours of work. *The Milbank Quarterly 85*(4), 641–690.

Maestas, N., M. Messel, and Y. Truskinovsky (2023). Caregiving and labor supply: New evidence from administrative data. Technical report, National Bureau of Economic Research.

Manser, M. and M. Brown (1980). Marriage and household decision-making: A bargaining analysis. *International Economic Review 21*(1), 31–44.

McElroy, M. B. and M. J. Horney (1981). Nash-bargained household decisions: Toward a generalization of the theory of demand. *International Economic Review 22*(2), 333–349.

Michalowsky, B., S. Flessa, T. Eichler, J. Hertel, A. Dreier, I. Zwingmann, D. Wucherer, H. Rau, J. R. Thyrian, and W. Hoffmann (2018). Healthcare utilization and costs in primary care patients with dementia: baseline results of the delphi-trial. *The European Journal of Health Economics 19*, 87–102.

Michalowsky, B., J. R. Thyrian, T. Eichler, J. Hertel, D. Wucherer, S. Flessa, and W. Hoffmann (2016). Economic analysis of formal care, informal care, and productivity losses in primary care patients who screened positive for dementia in germany. *Journal of Alzheimer's Disease 50*(1), 47–59.

Mommaerts, C. and Y. Truskinovsky (2020). The cyclicality of informal care. *Journal of health economics 71*, 102306.

Mommaerts, C. and Y. Truskinovsky (2023). Is all caregiving created equal? a comparison of caregiving to children and adults. In *AEA Papers and Proceedings*, Volume 113, pp. 627–631. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Nkangu, M. N., O. A. Olatunde, and S. Yaya (2017). The perspective of gender on the ebola virus using a risk management and population health framework: a scoping review. *Infectious Diseases of Poverty 6*(135).

Oreffice, S. and C. Quintana-Domeque (2020, July). Gender inequality in covid-19 times: Evidence from uk prolific participants. *IZA Discussion Paper* (13463).

Papageorge, N. W., M. V. Zahn, M. Belot, E. van den Broek-Altenburg, S. Choi, J. C. Jamison, and E. Tripodi (2020, June). Socio-demographic factors associated with self-protecting behavior during the covid-19 pandemic. Working Paper 27378, National Bureau of Economic Research.

Papanikolaou, D. and L. D. Schmidt (2020, June). Working remotely and the supply-side impact of covid-19. Working Paper 27330, National Bureau of Economic Research.

Ravindran, S. and M. Shah (2020, July). Unintended consequences of lockdowns: Covid-19 and the shadow pandemic. (27562).

Sevilla, A. and S. Smith (2020, May). Baby steps: The gender division of childcare during the covid-19 pandemic. *IZA Discussion Paper* (13302).

Skrip, L. A., M. P. Fallah, S. G. Gaffney, R. Yaari, D. Yamin, A. Huppert, L. Bawo, T. Nyenswah, and A. P. Galvani (2017). Characterizing risk of ebola transmission based on frequency and type of caseâcontact exposures. *Philosophical transactions of the Royal Society of London Series B, Biological sciences 372*(1721).

WHO (2007). Addressing sex and gender in epidemic-prone infectious diseases. *Departments of Gender, Women and Health, and Epidemic and Pandemic Alert and Response*.

Wiemers, E. E., S. Abrahams, M. AlFakhri, V. J. Hotz, R. F. Schoeni, and J. A. Seltzer (2020, June). Disparities in vulnerability to severe complications from covid-19 in the united states. Working Paper 27294, National Bureau of Economic Research.

World-Bank, a. (2019). *Mexico Gender Assessment*. World Bank.

Yi, J., J. J. Heckman, J. Zhang, and G. Conti (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal 125*(588), F347–F371.

Table 1: Labor Penalty in Mexico

	(1)	(2)
Panel A: Being Sick dummy		
Sick Partner	0.081***	0.076***
	(0.002)	(0.002)
Female	0.006***	-6.069
	(0.000)	(65.244)
Sick Partner $\times Female$	0.051***	0.047***
	(0.002)	(0.002)
Mean of Dependent Variable	0.012	0.012
Num.Obs.	2833092	2833092
Std.Errors	by: Household	by: Household
FE: Time	X	X
FE: Municipality	X	
FE: Individual		X
Panel B: Hours Worked Last Week		
Sick Partner	-5.203***	-4.225***
	(0.163)	(0.167)
Female	-10.994***	-2.557
	(0.030)	(11325.210)
Sick Partner $\times Female$	-0.880***	-0.506*
	(0.221)	(0.218)
Mean of Dependent Variable	39.397	39.397
Num.Obs.	2833092	2833092
Std.Errors	by: Household	by: Household
FE: Time	X	X
FE: Municipality	X	
FE: Individual		X

Notes: Being sick dummy is 1 if the respondent answered that they did not work last week and the reason was "illness, being excused, or personal affairs". Hours worked last week includes 0 if the respondent did not work at all last week. Sample includes working couples only. Robust standard errors are clustered at the household level and reported in parentheses.

Table 2: Labor Penalty in Mexico: Elderly Female

	(1)	(2)
Panel A: Being Sick dummy		
Sick Partner	0.069***	0.063***
	(0.002)	(0.003)
Female	-5.926	0.693
	(65.230)	(37.189)
Eld. Fem. Pres.	-0.002	0.000
	(0.002)	(0.005)
Sick Partner $\times Female$	0.054***	0.052***
	(0.002)	(0.010)
Sick Partner $\times Eld.Fem.Pres.$	0.074***	0.035*
	(0.010)	(0.016)
Female $\times Eld.Fem.Pres.$	0.006+	0.004
	(0.003)	(0.007)
Sick Partner $\times Female \times Eld.Fem.Pres.$	-0.062***	-0.005
	(0.018)	(0.032)
Mean of Dependent Variable	0.012	0.009
Num.Obs.	2833092	1878559
Std.Errors	by: Household	by: Household
FE: Time	X	X
FE: Individual	X	X
Sample	Full	Reweighted
Panel B: Hours Worked Last Week		
Sick Partner	-4.020***	-4.013***
	(0.174)	(0.289)
Female	-12.910	1.618
	(11325.037)	(4578.911)
Eld. Fem. Pres.	-0.863**	-0.401
	(0.265)	(0.649)
Sick Partner $\times Female$	-0.687**	-1.564+
	(0.227)	(0.890)
Sick Partner $\times Eld.Fem.Pres.$	-2.230***	-1.644
	(0.614)	(1.259)
Female $\times Eld.Fem.Pres.$	0.931*	-0.115
	(0.379)	(0.809)
Sick Partner $\times Female \times Eld.Fem.Pres.$	1.491	1.741
	(1.154)	(2.350)
Mean of Dependent Variable	39.397	41.107
Num.Obs.	2833092	1878559
Std.Errors	by: Household	by: Household
FE: Time	X	X
FE: Individual	X	X
Sample	Full	Reweighted

Notes: Being sick dummy is 1 if the respondent answered that they did not work last week and the reason was "illness, being excused, or personal affairs". Hours worked last week includes 0 if the respondent did not work at all last week. Robust standard errors are clustered at the household level and reported in parentheses.

Table 3: Labor Penalty in Mexico: Share of Female Income

	(1)	(2)
Panel A: Being Sick dummy		
Sick Partner	0.036***	0.034***
	(0.002)	(0.003)
Female	0.027***	-3.159
	(0.000)	(58.216)
Sh.fem.inc.	0.024***	0.026***
	(0.000)	(0.001)
Sick Partner $\times Female$	0.146***	0.137***
	(0.006)	(0.006)
Sick Partner $\times Sh.fem.inc.$	0.044***	0.039***
	(0.007)	(0.008)
Female $\times Sh.fem.inc.$	-0.051***	-0.052***
	(0.001)	(0.001)
Sick Partner $\times Female \times Sh.fem.inc.$	-0.205***	-0.188***
	(0.008)	(0.008)
Mean of Dependent Variable	0.011	0.011
Num.Obs.	2446167	2446167
Std.Errors	by: Household	by: Household
FE: Time	X	X
FE: Municipality	X	
FE: Individual		X
Panel B: Hours Worked Last Week		
Sick Partner	-5.258***	-3.685***
	(0.220)	(0.237)
Female	-20.921***	64.182
	(0.050)	(9904.889)
Sh.fem.inc.	-10.994***	-8.019***
	(0.077)	(0.086)
Sick Partner $\times Female$	-0.971*	-2.265***
	(0.446)	(0.456)
Sick Partner $\times Sh.fem.inc.$	1.425*	0.256
	(0.615)	(0.609)
Female $\times Sh.fem.inc.$	22.902***	14.159***
	(0.099)	(0.104)
Sick Partner $\times Female \times Sh.fem.inc.$	-1.723*	2.874***
	(0.781)	(0.772)
Mean of Dependent Variable	39.409	39.409
Num.Obs.	2446167	2446167
Std.Errors	by: Household	by: Household
FE: Time	X	X
FE: Municipality	X	
FE: Individual		X

Notes: Being sick dummy is 1 if the respondent answered that they did not work last week and the reason was "illness, being excused, or personal affairs". Hours worked last week includes 0 if the respondent did not work at all last week. Sample includes working couples only. Robust standard errors are clustered at the household level and reported in parentheses.

Table 4: Health Penalty: Unconditional and Conditional Infection Rate

	Member's early infection	Infection mean	Std.Dev.	Observations
All	0 or 1	0.0970	0.2959	80,478
Male	0 or 1	0.0887	0.2844	39,758
Female	0 or 1	0.1050	0.3066	40,720
All	0	0.0927	0.2899	70,944
Male	0	0.0847	0.2785	34,872
Female	0	0.1003	0.3004	36,072
All	1	0.1291	0.3353	9,534
Male	1	0.1175	0.3220	4,886
Female	1	0.1414	0.3484	4,648
All	Male	0.1368	0.3437	4,606
Male	Male	0.0864	0.2812	602
Female	Male	0.1444	0.3515	4,004
All	Female	0.1223	0.3277	5,625
Male	Female	0.1208	0.3259	4,594
Female	Female	0.1290	0.3354	1,031

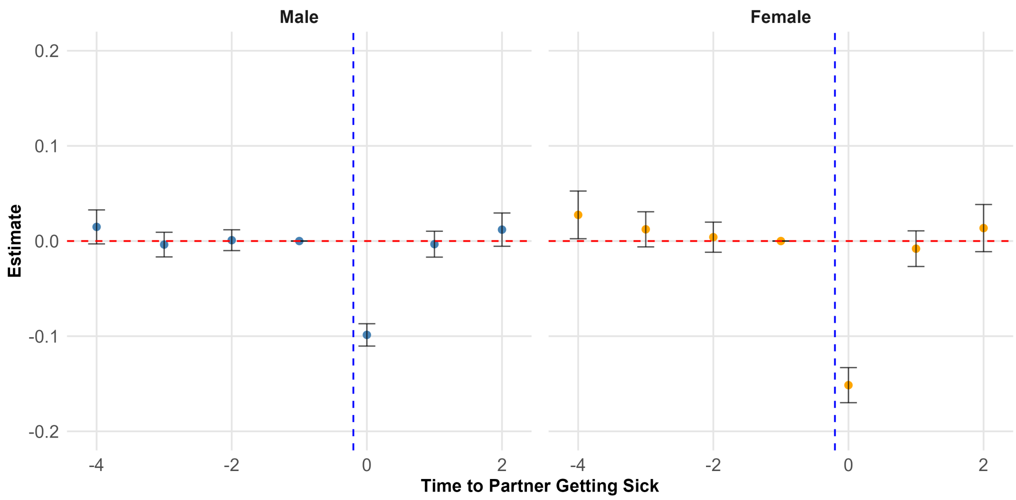
Notes: There are 3909 individuals with only male member's early infection, 4928 only female's, 697 with both male and female's early infection

Table 5: Household Transmission of Infectious Disease

	Infection dummy			
	(1)	(2)	(3)	(4)
Member	1.281	1.292	1.275	1.198
	(0.631)	(0.631)	(0.630)	(0.638)
Female	1.695	1.691	1.684	1.679
	(0.270)	(0.271)	(0.271)	(0.272)
Member×Female	1.049	1.055	1.043	1.032
	(0.491)	(0.490)	(0.491)	(0.492)
Age	-0.536	-0.533	-0.516	-0.517
	(0.064)	(0.064)	(0.065)	(0.061)
Age^2	0.006	0.006	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.081		
		(0.067)		
Income				-0.001
				(0.004)
Home value				-0.003
				(0.001)
R^2	0.0352	0.0352	0.0353	0.0357
Observations	80478	80478	80478	80478
Y-mean	9.697	9.697	9.697	9.697
Zip FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Figure 1: Poisson Event Study: Working Hours



Note: The figure shows event study estimates of the impact of a partner's illness on working hours using Poisson model, separated by gender. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue precedes the interview time with the event

For Online Publication

S1 Additional Assumptions in Conceptual Frameworks

S1.1 Family bargaining

In the main conceptual framework, we assume two family members are in total agreement, and they jointly conduct household consumption and jointly maximize their household utility. Now we assume members negotiate with each other to reach an agreement compromising different individual preferences. Based on Manser and Brown (1980) and McElroy and Horney (1981), we use the cooperative approach according to the Nash solution for further analysis. If they do not reach an agreement, they will instead choose their outside options. The Nash solution is characterized by cooperative gain sharing in order to maximize the product of the two individual gains:

$$max\ U_{C_m,C_f,x_m} = [U_m(x_m) - D_m]^b \cdot [U_f(x_f) - D_f]^{1-b}$$

where x_m, x_f is male's and female's private consumption good. D_m, D_f is male's and female's payoff when the partner does not agree, namely his or her outside option. The male member's bargaining power is an exogenous parameter b. C_m, C_f, budget constraint and health production functions are the same as those in the main model.

$$x_m + x_f \leq w_m(T - C_m) + w_f(T - C_f) + P(H_m - \bar{H}_0) + P(H_f - \bar{H}_0)$$

$$H_m = \bar{H}_0 + \alpha lnC_f$$

$$H_f = \bar{H}_0 + \beta lnC_m$$

The first-order conditions generate:

$$max\ \mathcal{L}_{C_m,C_f,x_m} = [U_m(x_m) - D_m]^b \cdot [U_f(x_f) - D_f]^{1-b}$$

$$+\lambda[w_m(T - C_m) + w_f(T - C_f) + P\alpha lnC_f + P\beta lnC_m - x_m - x_f]$$

$$\left.\frac{\partial \mathcal{L}}{\partial C_m}\right|_{C_m^*} = \lambda(-w_m + \frac{P\beta}{C_m^*}) = 0$$

$$\left.\frac{\partial \mathcal{L}}{\partial C_f}\right|_{C_f^*} = \lambda(-w_f + \frac{P\alpha}{C_f^*}) = 0$$

$$\frac{C_f^*}{C_m^*} = \frac{\alpha}{\beta} \cdot \frac{w_m}{w_f}$$

$$H_m^* - H_f^* = (\alpha ln\alpha - \beta ln\beta) + (\alpha - \beta) \cdot lnP - \alpha lnw_f + \beta lnw_m$$

The optimal level of each member's care and health condition is exactly the same as the main results without family bargaining. This results from the assumption that health is a pure investment good and does not enter the utility function, the only component affected by the bargaining assumption.

S1.2 Too costly to provide care

Under the condition of male's initial infection of an infectious disease, now we assume $0 < \alpha \leq \gamma < 1$. The marginal cost of providing care (on the care provider's health) is higher than the marginal return of health (on the care receiver's health). In other words, transmission of the disease is very strong, so it's very risky to provide care to the infected member. Solving the first-order condition, household utility is decreasing with female's caregiving time, so she will provide the minimum level of care:

$$\frac{\partial x_c}{\partial C_f} = -w_f + \frac{P\alpha}{C_f} - \frac{P\gamma}{C_f} < 0$$

$$\frac{\partial x_c}{\partial C_m} = -w_m + \frac{P\beta}{C_m}$$

$$C_f^{M*} = 1 < C_f^* = \frac{P\alpha}{w_f}$$

$$C_m^{S*} = C_m^* = \frac{P\beta}{w_m}$$

$$H_f^{M*} = H_f^*$$

$$H_m^{S*} = \bar{H}_0 - S < H_m^* - S$$

There is no care allocated to the infected male member, and no household penalty for female. The male member is even worse than his condition under infection-free case by a level larger than S, the shock itself.

Under this extreme and brutal condition, family members will give up on the infected member and leave him/her aside. We should not see any family cluster of infectious diseases, which is contradicted with our observations. Household transmission aside, the strong infectious diseases must be controlled in a severe way like quarantine in a hospital because contact in everyday life is also risky. This is less likely to happen in the common infectious disease, so we focus on conditions with $0 < \gamma < \alpha < 1$ in the main conceptual framework.

S2 Impact of Sick Children

We examine the impact of children's health shocks on the labor supply responses of parents, utilizing data from the 2012 and 2018 waves of the National Health and Nutrition Survey (ENSANUT) in Mexico. ENSANUT provides a representative cross-sectional survey encompassing 50,000 households, with detailed information on health, nutrition, and labor market outcomes. We restrict our sample following the same requirements as for the ENOE, and to households with at least one child, which results in a sample of approximately 40,000 employed adults.

Our primary explanatory variable is the incidence of child illness within the past two weeks, which is further disaggregated by the child's age to assess whether the effects are more pronounced for younger children. This variable is derived from a question regarding householder's health problems in the the last two weeks. The dependent variable of interest is the number of hours worked in the past week by an employed adult, with a mean of 46.6 hours for men and 38.4 hours for women.

We estimate the following model to identify the labor supply response to child health shocks:

$$\text{Hours Worked Last Week}_i = \beta_0 + \beta_1 \text{Any child sick in last 2 weeks}_i + \beta_2 \text{Female}_i \tag{S.1}$$

$$+ \beta_3 \text{Any child sick in last 2 weeks}_i \times \text{Female}_i + \beta_4 X_i + \epsilon_i \tag{S.2}$$

The coefficients of primary interest are β_1 and β_3. β_1 captures the change in hours worked by men following a child's illness, while β_3 represents the differential response of women's labor supply relative to men's in the event of a child's health shock. Thus it indicates the additional penalty borne by women. We further extend our analysis by disaggregating the illness variable by child age. The vector X_i includes controls for individual and household incomes, age dummies, number of children, education levels, and survey timing. Despite the comprehensive set of controls, the cross-sectional nature of our data requires caution in causal interpretation, as unobserved characteristics may differ between parents of sick and healthy children.

Our findings, presented in Table S4, reveal significant gender differences in labor supply responses to child illness. Column (1) shows that men slightly increase their working hours by approximately 0.5 hours (1% of average male working hours) in response to child illness. In contrast, women reduce their labor supply by 2.75 hours (7% of average female weekly hours).

These results suggest that women assume a larger share of caregiving responsibilities during child health shocks. Column (2) further indicates that the reduction in female labor supply is more pronounced for younger children, with a decrease of 3.3 hours for children under 10 years old, compared to a 1.9-hour reduction for children aged 11 to 19. The age of the child has a less significant impact on fathers' labor supply responses. Overall, our results demonstrate that child illness exerts a considerably stronger effect on mothers' labor supply compared to fathers'.

S3 Differential Health-Seeking Behavior

In this section, we address the potential concern that differential health-seeking behavior between men and women might influence our results. Specifically, we demonstrate that the observed gender disparity in infection rates cannot be entirely driven by differences in the propensity to seek medical attention.

To formalize this, let Y_{ih}^* represent the true health status of individual i in household h. For men, we model this as a function of constant gender characteristics α, and a treatment effect γ of having an infected household member, and an error term ε_{ih}:

$$Y_{ih}^* = \alpha + \gamma Member_{ih} + \varepsilon_{ih} \tag{S.3}$$

Since our data are derived from insurance claims, we observe Y_{ih}, the reported health status, which is the true health status scaled by the propensity to seek medical attention δ_m for men:

$$Y_{ih} = \delta_m Y_{ih}^* = \delta_m \alpha + \delta_m \gamma Member_{ih} + \delta_m \varepsilon_{ih} \tag{S.4}$$

For women, the true health status is modeled similarly with potentially different intercept ψ and treatment effect ζ:

$$Y_{ih}^* = \psi + \zeta Member_{ih} + \varepsilon_{ih} \tag{S.5}$$

The reported health status for women is scaled by their propensity to seek medical attention δ_f:

$$Y_{ih} = \delta_f Y_{ih}^* = \delta_f \psi + \delta_f \zeta Member_{ih} + \delta_f \varepsilon_{ih} \tag{S.6}$$

Our empirical estimates capture the observed infection rates for both genders. The model we estimate is:

$$Y_{ih} = \underbrace{\delta_m \alpha}_{\beta_0} + \underbrace{\delta_m \gamma}_{\beta_1} Member_{ih} + \underbrace{(\delta_f \psi - \delta_m \alpha)}_{\beta_2} Female_i + \underbrace{(\delta_f \zeta - \delta_m \gamma)}_{\beta_3} Female_i \times Member_{ih} + e_{ih} \tag{S.7}$$

Here, β_0 is the baseline reported infection probability for men, β_1 captures the effect of household member infection for men on reported infections, β_2 represents the difference in baseline infection report probabilities between women and men, and β_3 measures the differential impact of household member infection on reporting of women relative to men.

The true female household penalty is $\zeta - \gamma$. With differential reporting, we observe $(\delta_f \zeta - \delta_m \gamma)$. If the observed gender disparity were solely driven by differential propensity to visit the doctor (and $\zeta = \gamma$), the ratio of the male coefficient to the interaction coefficient $\frac{\beta_3 - \beta_1}{\beta_1}$ should equal $\frac{\delta_f - \delta_m}{\delta_m}$, representing the change in probability for female reporting.

In the claims data, we do not observe the propensity to report an infection to a doctor. Hence, to estimate the quantity $\frac{\delta_f - \delta_m}{\delta_m}$, we leverage a unique survey data from ENSANUT, which asks both about health problems and whether they were reported to a doctor. ENSANUT is a representative health survey collected in Mexico. We use wave from the year 2012. It provides information on whether individuals visited a doctor conditional on having a health problem in the last two weeks. The survey includes variables capturing visits to the doctor under different circumstances: any visit to the doctor, visiting the doctor for non-serious conditions, and visiting the doctor for serious conditions. To capture δ_g, the propensity to visit a doctor, we regress whether an individual with a health problem visited a doctor on their gender:

$$y_i = \lambda_0 + \lambda_1 Female_i + u_i \tag{S.8}$$

The results are presented in Table S1. In this context λ_0 is the propensity to visit a doctor for men and λ_1 is the difference between female and male propensity. Consequently, according to the first specification, $\frac{\delta_f - \delta_m}{\delta_m} = \frac{\lambda_1}{\lambda_0} \approx 0.19$. This implies that differences in reporting propensities could only produce a female coefficient that is 19% larger than the male coefficient, whereas we observe much larger differences in our data (of around 80%). Therefore, differential health-seeking behavior can only explain a small part of the observed gender disparity, and the substantial differences we observe are likely due to underlying differences in infection patterns.

Table S1: Propensity to Visit Doctors by Gender

Dependent Variables:	Visit Doctor	Visit if Not Serious	Visit if Serious
Model:	(1)	(2)	(3)
Variables			
Constant	0.5195***	0.4104***	0.7261***
	(0.0074)	(0.0106)	(0.0137)
Female	0.1013***	0.0902***	0.0747***
	(0.0095)	(0.0139)	(0.0174)
Fit statistics			
Observations	11,023	5,267	2,452
Adjusted R^2	0.01004	0.00775	0.00707

Notes: IID standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

S4 Additional Heterogeneity Results

S4.1 Heterogeneous effects by race composition

We test the heterogeneous effects using zipcode-level race composition. Table S18 shows how the proportion of the black group affects the household penalty. Areas with larger proportions of minority groups have lower household transmission rate, and its gender difference is not significantly different from that in areas with fewer minority group. This may result from the baseline worse health conditions of the minority group in the control group. If the minority group faces higher contagion outside households, it may absorb the observed household transmission effects.

S4.2 Heterogeneous effects by jobs

Jobs are roughly classified in our sample, including real estate, sales, government, educator, etc. Given this constraint, we are not able to explore heterogeneity by jobs based on the level of inter-personal interactions or the capacity of working from home. In this section, we only test if being a healthcare worker affects the infection rate. To do so, we add an additional dummy variable to control for the individual being a healthcare worker. In Table S19 Panel A, the imprecise estimate on $Healthcare$ suggests no significant impact of being a healthcare worker on his/her own infection rate, though the sign is positive. The coefficients of interest remain stable with the additional control. Furthermore, we test if being a healthcare worker or if there is a healthcare worker in the household affects β_1 and β_3 using a triple-difference design. Results in Table S19 Panel B and C show it affects neither the household penalty nor its gender difference. This concludes that medical workers may face slightly higher risk of infection outside the household, but intra-household transmission pattern is not significantly different from that faced by people in other occupations.

S5 Additional Tables

Table S2: Labor Penalty: Reweighted Sample

	(1)	(2)
Panel A: Being Sick dummy		
Sick Partner	0.016***	0.043***
	(0.003)	(0.003)
Female	0.036***	-0.331
	(0.001)	(48.667)
Sick Partner $\times Female$	0.014***	0.046***
	(0.002)	(0.002)
Mean of Dependent Variable	0.009	0.009
Num.Obs.	1878559	1878559
Std.Errors	by: Household	by: Household
FE: Time	X	X
FE: Municipality	X	
FE: Individual		X
Panel B: Hours Worked Last Week		
Sick Partner	-2.034***	-3.716***
	(0.208)	(0.205)
Female	-11.051***	-2.183
	(0.087)	(3110.932)
Sick Partner $\times Female$	0.905**	-0.373
	(0.281)	(0.270)
Mean of Dependent Variable	41.107	41.107
Num.Obs.	1878559	1878559
Std.Errors	by: Household	by: Household
FE: Time	X	X
FE: Municipality	X	
FE: Individual		X

Notes: We use variables from the balance table S7 to construct the propensity score of having a sick household member and apply these scores to re-weight the sample, ensuring it is balanced on observable characteristics. The "Being sick" dummy variable equals 1 if the respondent reported not working last week due to "illness, being excused, or personal affairs." "Hours worked last week" includes a value of 0 if the respondent did not work at all during that week. The sample includes only working couples. Robust standard errors are clustered at the household level and are reported in parentheses.

Table S3: Balance Table: Elderly presence

	HH without Elderly Female	HH with Elderly Female	Difference
Household Size	3.23 (1.2)	3.91 (1.4)	0.68 (0.01)
Yearly Income (USD)	7822.19 (8140.5)	6859.17 (8658.2)	-963.02 (60.4)
Age	41.71 (10.3)	52.66 (13.9)	10.95 (0.092)
Female	0.51 (0.5)	0.28 (0.45)	-0.23 (0.002)
Working Last Week	0.96 (0.207)	0.95 (0.212)	-0.00 (0.001)
Hours Worked Last Week	42.26 (17.735)	43.20 (18.281)	0.94 (0.1)
Female Income Share	0.42 (0.28)	0.43 (0.3)	0.02 (0.002)

Notes: Data comes from the Mexican Labor Survey. The statistics are calculated for the first period of interviews, during which all households were initially healthy. Standard errors are provided in parentheses.

Table S4: Children's Sickness and Parents Labor Outcomes

Dependent Variable:	Hours worked last week	
Model:	(1)	(2)
Any child sick in last 2 weeks	0.5486	
	(0.2795)	
Female	-10.16	-10.30
	(0.2275)	(0.2250)
Any child sick in last 2 weeks × Female	-3.338	
	(0.4895)	
Sick child aged 0-1 in last 2 weeks		0.8343
		(0.6273)
Sick child aged 2-10 in last 2 weeks		0.8074
		(0.3887)
Sick child aged 11-19 in last 2 weeks		0.5359
		(0.5014)
Sick child aged 0-1 in last 2 weeks × Female		-3.851
		(1.148)
Sick child aged 2-10 in last 2 weeks × Female		-4.227
		(0.6767)
Sick child aged 11-19 in last 2 weeks × Female		-2.373
		(0.7979)
Fixed-effects		
Age, Education, Survey Time	Yes	Yes
Total number of children	Yes	
Number of children aged 0-1		Yes
Number of children aged 2-10		Yes
Number of children aged 11-19		Yes
Adjusted R^2	0.07768	0.07825
Observations	40,375	40,375
Dependent variable mean	44.680	44.680

Notes: Data comes from the National Health and Nutrition Survey in Mexico (ENSANUT) rounds 2012 and 2018. Standard errors are clustered at the household level.

Table S5: Summary Statistics in Insurance Claims

	Observations	Mean	Std.Dev.	Min	Max
Infection	80,478	0.0970	0.2959	0	1
Member's early infection	80,478	0.1185	0.3232	0	1
Female	80,478	0.5060	0.5000	0	1
Age	80,478	59.16	15.70	19	80
Household size	80,478	4.495	1.827	2	10
#Children in household	80,478	1.271	1.133	0	4
#Adults in household	80,478	3.224	1.198	1	6

Table S6: Summary Statistics: Full ENOE Sample

Variable	Mean	StdDev	Min	Max
Sick	0.012	0.109	0	1.0
Sick Partner	0.012	0.109	0	1.0
Female	0.500	0.500	0	1.0
Age	42.747	10.709	19	80.0
Household Size	3.302	1.265	2	10.0
Household income (USD)	7569.588	7945.456	0	653635.3
Female Work Hours	33.905	19.030	0	144.0
Male Work Hours	44.882	19.202	0	168.0
Female Worked Last Week	0.933	0.250	0	1.0
Male Worked Last Week	0.949	0.219	0	1.0

Table S7: Balance Table: ENOE

	Control	Treated	Difference
Household Size	3.27 (1.261)	3.18 (1.257)	-0.09 (0.009)
Yearly Income (USD)	7793.63 (8217.673)	7429.02 (7378.392)	-364.61 (56.096)
Age	42.25 (10.750)	42.78 (11.455)	0.53 (0.083)
Female	0.50 (0.500)	0.50 (0.500)	0.00 (0.000)
Working Last Week	0.96 (0.206)	0.94 (0.231)	-0.01 (0.001)
Hours Worked Last Week	42.38 (17.736)	41.20 (18.189)	-1.17 (0.102)
Female Income Share	0.42 (0.290)	0.43 (0.293)	0.01 (0.002)

Notes: Data comes from the Mexican Labor Survey. The statistics are calculated for the first period of interviews, during which all households were initially healthy. The "Treated" column represents households where at least one member became sick during the interview periods. The "Control" column represents households where all members remained healthy throughout. Standard errors are provided in parentheses.

Table S8: Household Transmission of Infectious Disease, Children and Adults in Household Added

	Infection dummy		
	(1)	(2)	(3)
Member	1.296	1.272	1.195
	(0.631)	(0.628)	(0.635)
Female	1.696	1.687	1.682
	(0.271)	(0.272)	(0.273)
Member×Female	1.066	1.057	1.044
	(0.491)	(0.493)	(0.493)
Age	-0.548	-0.529	-0.527
	(0.064)	(0.064)	(0.061)
Age^2	0.006	0.005	0.005
	(0.001)	(0.001)	(0.001)
#Children	0.090		
in household	(0.107)		
#Adults	-0.223		
in household	(0.090)		
Income			-0.000
			(0.004)
Home value			-0.003
			(0.001)
R^2	0.0352	0.0355	0.0359
Observations	80478	80478	80478
Y-mean	9.697	9.697	9.697
ZIP FEs	Y	Y	Y
#Children FEs		Y	Y
#Adults FEs		Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S9: Household Transmission of Infectious Disease, 1-4 Years

	Panel A: Infection 1-2 years			
	(1)	(2)	(3)	(4)
Member	0.205	0.216	0.197	0.119
	(0.544)	(0.545)	(0.543)	(0.549)
Female	1.694	1.690	1.682	1.676
	(0.270)	(0.270)	(0.271)	(0.272)
Member×Female	0.579	0.585	0.573	0.563
	(0.468)	(0.467)	(0.468)	(0.469)
Age	-0.542	-0.540	-0.522	-0.520
	(0.060)	(0.061)	(0.061)	(0.058)
Age^2	0.006	0.006	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.081		
		(0.066)		
Income				-0.002
				(0.004)
Home value				-0.002
				(0.001)
R^2	0.0322	0.0322	0.0324	0.0328
Observations	80478	80478	80478	80478
Y-mean	9.532	9.532	9.532	9.532
	Panel B: Infection 2-3 years			
Member	0.026	0.038	0.019	-0.059
	(0.575)	(0.575)	(0.574)	(0.581)
Female	1.691	1.687	1.679	1.673
	(0.270)	(0.270)	(0.271)	(0.271)
Member×Female	0.508	0.514	0.503	0.492
	(0.512)	(0.512)	(0.512)	(0.512)
Age	-0.537	-0.535	-0.517	-0.516
	(0.061)	(0.062)	(0.062)	(0.058)
Age^2	0.006	0.006	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.083		
		(0.066)		
Income				-0.001
				(0.004)
Home value				-0.002
				(0.001)
R^2	0.0318	0.0319	0.0320	0.0324
Observations	80478	80478	80478	80478
Y-mean	9.506	9.506	9.506	9.506
	Panel C: Infection 3-4 years			
Member	-0.168	-0.157	-0.175	-0.254
	(0.589)	(0.590)	(0.588)	(0.595)
Female	1.690	1.686	1.678	1.673

	(0.270)	(0.270)	(0.270)	(0.271)
Member×Female	0.474	0.480	0.468	0.458
	(0.489)	(0.489)	(0.489)	(0.490)
Age	-0.538	-0.535	-0.517	-0.516
	(0.060)	(0.061)	(0.062)	(0.058)
Age^2	0.006	0.006	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.083		
		(0.065)		
Income				-0.002
				(0.004)
Home value				-0.002
				(0.001)
R^2	0.0316	0.0317	0.0318	0.0322
Observations	80478	80478	80478	80478
Y-mean	9.480	9.480	9.480	9.480
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S10: Within Household Comparison

	Infection dummy
Female	2.745
	(1.272)
Age	-0.8123
	(0.3552)
Age^2	0.0085
	(0.0034)
R^2	0.7928
Observations	3061
Y-mean	14.83
Household FEs	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the household level, reported in parentheses.

Table S11: Household Transmission of Infectious Disease after Male's Infection

	Infection dummy			
	(1)	(2)	(3)	(4)
Member_m	-0.340	-0.251	-0.292	-0.461
	(1.188)	(1.190)	(1.188)	(1.178)
Female	1.656	1.651	1.636	1.636
	(0.267)	(0.267)	(0.267)	(0.269)
$\text{Member}_m \times \text{Female}$	2.758	2.681	2.689	2.791
	(1.373)	(1.376)	(1.373)	(1.366)
Age	-0.523	-0.520	-0.502	-0.499
	(0.066)	(0.067)	(0.067)	(0.063)
Age^2	0.005	0.005	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.093		
		(0.068)		
Income				-0.002
				(0.004)
Home value				-0.003
				(0.001)
R^2	0.0348	0.0349	0.0350	0.0356
Observations	75550	75550	75550	75550
Y-mean	9.741	9.741	9.741	9.741
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: The smaller sample size is because 4928 individuals with only female member's early infection are dropped. Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S12: Household Transmission of Infectious Disease after Female's Infection

	Infection dummy			
	(1)	(2)	(3)	(4)
Member_f	1.585	1.592	1.574	1.503
	(0.661)	(0.661)	(0.660)	(0.666)
Female	1.703	1.698	1.690	1.686
	(0.270)	(0.270)	(0.270)	(0.271)
$\text{Member}_f \times \text{Female}$	1.089	1.155	1.140	0.983
	(1.283)	(1.278)	(1.283)	(1.293)
Age	-0.525	-0.522	-0.503	-0.504
	(0.067)	(0.068)	(0.069)	(0.066)
Age^2	0.005	0.005	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.084		
		(0.068)		
Income				-0.000
				(0.004)
Home value				-0.003
				(0.001)
R^2	0.0330	0.0330	0.0332	0.0336
Observations	76569	76569	76569	76569
Y-mean	9.736	9.736	9.736	9.736
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: The smaller sample size is because 3909 individuals with only male member's early infection are dropped. Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S13: Household Transmission of Infectious Disease by the First Patient's Gender

	Infection dummy			
	(1)	(2)	(3)	(4)
Member_m	-0.238	-0.156	-0.209	-0.340
	(1.290)	(1.288)	(1.286)	(1.279)
Member_f	3.051	3.054	3.040	2.975
	(0.659)	(0.658)	(0.656)	(0.662)
Female	1.699	1.694	1.688	1.683
	(0.271)	(0.271)	(0.272)	(0.273)
$\text{Member}_m \times \text{Female}$	4.198	4.119	4.143	4.235
	(1.600)	(1.601)	(1.601)	(1.596)
$\text{Member}_f \times \text{Female}$	-0.369	-0.297	-0.306	-0.454
	(1.369)	(1.366)	(1.369)	(1.373)
Age	-0.530	-0.527	-0.510	-0.512
	(0.063)	(0.064)	(0.064)	(0.061)
Age^2	0.006	0.005	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.088		
		(0.065)		
Income				-0.000
				(0.004)
Home value				-0.003
				(0.001)
R^2	0.0362	0.0362	0.0364	0.0368
Observations	80478	80478	80478	80478
Y-mean	9.697	9.697	9.697	9.697
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S14: Household Transmission of Infectious Disease by the First Patient's Age

	Infection dummy			
	(1)	(2)	(3)	(4)
Member_{old}	4.403	4.406	4.379	4.323
	(0.799)	(0.799)	(0.797)	(0.801)
Member_{middle}	0.773	0.784	0.799	0.705
	(0.936)	(0.935)	(0.934)	(0.947)
Member_{young}	3.591	3.638	3.596	3.476
	(1.132)	(1.132)	(1.138)	(1.136)
Female	1.688	1.683	1.677	1.673
	(0.273)	(0.273)	(0.273)	(0.274)
$\text{Member}_{old}\times$Female	0.888	0.894	0.877	0.883
	(0.617)	(0.618)	(0.618)	(0.617)
$\text{Member}_{middle}\times$Female	2.057	2.060	2.059	2.008
	(1.011)	(1.011)	(1.013)	(1.015)
$\text{Member}_{young}\times$Female	-0.686	-0.666	-0.652	-0.644
	(1.589)	(1.586)	(1.588)	(1.588)
Age	-0.519	-0.516	-0.501	-0.504
	(0.062)	(0.063)	(0.064)	(0.060)
Age^2	0.005	0.005	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.089		
		(0.066)		
Income				-0.000
				(0.004)
Home value				-0.003
				(0.001)
R^2	0.0367	0.0368	0.0369	0.0373
Observations	80478	80478	80478	80478
Y-mean	9.697	9.697	9.697	9.697
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S15: Heterogenous Effects by Member's Age

	Infection dummy		
	Young	Middle	Old
	(1)	(2)	(3)
Member_{old}	0.962	0.611	4.994
	(2.314)	(1.414)	(0.898)
Member_{middle}	-1.277	-0.425	2.675
	(1.833)	(1.051)	(1.661)
Member_{young}	0.823	4.863	1.900
	(2.025)	(1.806)	(2.196)
Female	2.478	1.760	1.352
	(0.621)	(0.471)	(0.297)
$\text{Member}_{old}\times\text{Female}$	1.830	4.849	0.588
	(3.004)	(1.826)	(0.690)
$\text{Member}_{middle}\times\text{Female}$	10.885	0.787	-1.688
	(2.697)	(0.944)	(2.180)
$\text{Member}_{young}\times\text{Female}$	1.082	-2.159	-3.947
	(2.389)	(2.332)	(3.196)
Age	-0.912	0.493	-2.211
	(0.725)	(0.555)	(0.515)
Age^2	0.014	-0.005	0.018
	(0.012)	(0.005)	(0.004)
Income	0.005	-0.004	0.005
	(0.011)	(0.005)	(0.006)
Home value	-0.005	-0.003	-0.003
	(0.002)	(0.001)	(0.001)
R^2	0.0719	0.0447	0.0451
Observations	10302	29455	40721
Y-mean	9.913	8.299	10.61
ZIP FEs	Y	Y	Y
Household size FEs	Y	Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S16: Heterogenous Effects by Zipcode-level Gender Income Gap

	Infection dummy			
	Panel A: Areas where gender income gap is below the median			
	(1)	(2)	(3)	(4)
Member	1.433	1.369	1.353	1.267
	(1.183)	(1.188)	(1.193)	(1.197)
Female	1.857	1.870	1.859	1.858
	(0.456)	(0.457)	(0.458)	(0.455)
Member×Female	0.024	0.007	-0.037	-0.014
	(1.307)	(1.302)	(1.310)	(1.309)
Age	-0.769	-0.778	-0.754	-0.754
	(0.103)	(0.103)	(0.107)	(0.108)
Age^2	0.008	0.008	0.008	0.008
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		0.295		
		(0.116)		
Income				-0.005
				(0.008)
Home value				-0.002
				(0.002)
R^2	0.0505	0.0507	0.0511	0.0514
Observations	21127	21127	21127	21127
Y-mean	10.37	10.37	10.37	10.37
	Panel B: Areas where gender income gap is above the median			
Member	1.269	1.290	1.270	1.201
	(0.652)	(0.651)	(0.653)	(0.655)
Female	1.646	1.635	1.627	1.622
	(0.339)	(0.339)	(0.339)	(0.339)
Member×Female	1.244	1.260	1.257	1.235
	(0.672)	(0.672)	(0.672)	(0.672)
Age	-0.454	-0.448	-0.434	-0.445
	(0.080)	(0.081)	(0.080)	(0.079)
Age^2	0.005	0.005	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.185		
		(0.068)		
Income				0.003
				(0.005)
Home value				-0.003
				(0.001)
R^2	0.0280	0.0281	0.0282	0.0286
Observations	56392	56392	56392	56392
Y-mean	9.468	9.468	9.468	9.468
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S17: Heterogenous Effects by Zipcode-level Time Use Difference

	Infection dummy			
	Panel A: Areas where home production gap is below the median			
	(1)	(2)	(3)	(4)
Member	2.632	2.633	2.613	2.514
	(1.027)	(1.027)	(1.030)	(1.042)
Female	1.418	1.417	1.414	1.407
	(0.454)	(0.455)	(0.456)	(0.450)
Member×Female	0.357	0.358	0.332	0.310
	(0.848)	(0.848)	(0.848)	(0.851)
Age	-0.718	-0.717	-0.698	-0.689
	(0.121)	(0.121)	(0.119)	(0.120)
Age^2	0.007	0.007	0.007	0.007
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.011		
		(0.093)		
Income				-0.003
				(0.006)
Home value				-0.003
				(0.002)
R^2	0.0254	0.0254	0.0256	0.0264
Observations	27528	27528	27528	27528
Y-mean	9.405	9.405	9.405	9.405
	Panel B: Areas where home production gap is above the median			
Member	0.862	0.868	0.877	0.794
	(1.055)	(1.053)	(1.055)	(1.052)
Female	1.561	1.555	1.549	1.539
	(0.550)	(0.548)	(0.549)	(0.551)
Member×Female	2.233	2.238	2.213	2.213
	(1.211)	(1.214)	(1.214)	(1.215)
Age	-0.526	-0.524	-0.506	-0.523
	(0.102)	(0.102)	(0.104)	(0.103)
Age^2	0.005	0.005	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.064		
		(0.120)		
Income				0.007
				(0.008)
Home value				-0.005
				(0.002)
R^2	0.0415	0.0416	0.0418	0.0423
Observations	23141	23141	23141	23141
Y-mean	9.874	9.874	9.874	9.874
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S18: Heterogenous Effects by Race Composition

	Infection dummy			
	(1)	(2)	(3)	(4)
Member	2.438	2.445	2.430	2.376
	(0.510)	(0.511)	(0.513)	(0.515)
Female	1.725	1.722	1.716	1.717
	(0.223)	(0.224)	(0.221)	(0.223)
Member×Female	0.924	0.930	0.921	0.902
	(0.458)	(0.456)	(0.462)	(0.463)
Black×Member	-0.137	-0.137	-0.138	-0.140
	(0.049)	(0.049)	(0.050)	(0.049)
Black×Female	-0.002	-0.002	-0.002	-0.003
	(0.018)	(0.017)	(0.017)	(0.018)
Black×Member×Female	-0.002	-0.002	-0.002	-0.001
	(0.057)	(0.057)	(0.058)	(0.058)
Age	-0.535	-0.533	-0.517	-0.522
	(0.055)	(0.054)	(0.059)	(0.057)
Age^2	0.006	0.006	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.058		
		(0.036)		
Income				-0.001
				(0.002)
Home value				-0.003
				(0.001)
R^2	0.0347	0.0347	0.0348	0.0352
Observations	80478	80478	80478	80478
Y-mean	9.697	9.697	9.697	9.697
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Infection dummy is multiplied by 100. Black is in percentage and is absorbed by zipcode fixed effects. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S19: Healthcare Workers

	Infection dummy			
	Panel A: Additional control			
	(1)	(2)	(3)	(4)
Member	1.281	1.292	1.275	1.196
	(0.630)	(0.630)	(0.629)	(0.636)
Female	1.689	1.684	1.675	1.657
	(0.268)	(0.268)	(0.269)	(0.270)
Member×Female	1.050	1.056	1.044	1.035
	(0.491)	(0.491)	(0.492)	(0.492)
Healthcare	0.098	0.115	0.144	0.397
	(0.520)	(0.519)	(0.520)	(0.520)
Age	-0.537	-0.535	-0.518	-0.521
	(0.066)	(0.067)	(0.067)	(0.063)
Age^2	0.006	0.006	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.081		
		(0.066)		
Income				-0.001
				(0.004)
Home value				-0.003
				(0.001)
R^2	0.0352	0.0352	0.0353	0.0358
Observations	80478	80478	80478	80478
Y-mean	9.697	9.697	9.697	9.697
	Panel B: Triple difference			
Member	1.287	1.298	1.281	1.203
	(0.617)	(0.617)	(0.616)	(0.620)
Female	1.658	1.652	1.644	1.629
	(0.269)	(0.269)	(0.269)	(0.271)
Member×Female	0.944	0.950	0.938	0.931
	(0.507)	(0.507)	(0.508)	(0.509)
Healthcare	-0.998	-0.987	-0.947	-0.583
	(0.934)	(0.935)	(0.931)	(0.950)
Healthcare×Member	-0.155	-0.152	-0.167	-0.242
	(3.283)	(3.284)	(3.301)	(3.367)
Healthcare×Female	1.237	1.243	1.231	1.097
	(1.232)	(1.231)	(1.230)	(1.248)
Healthcare×Member×Female	1.604	1.609	1.624	1.644
	(3.370)	(3.373)	(3.393)	(3.446)
Age	-0.538	-0.535	-0.518	-0.522
	(0.066)	(0.067)	(0.067)	(0.063)
Age^2	0.006	0.006	0.005	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.082		
		(0.066)		

Income				-0.001
				(0.004)
Home value				-0.003
				(0.001)
R^2	0.0352	0.0352	0.0354	0.0358
Observations	80478	80478	80478	80478
Y-mean	9.697	9.697	9.697	9.697
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S20: Results with Alternative Sample

	Infection dummy			
	Panel A: Require two individuals in households			
	(1)	(2)	(3)	(4)
Member	1.472	1.509	1.500	1.521
	(0.639)	(0.636)	(0.636)	(0.639)
Female	1.740	1.746	1.736	1.683
	(0.193)	(0.193)	(0.193)	(0.197)
Member×Female	0.949	0.947	0.941	0.986
	(0.539)	(0.540)	(0.540)	(0.539)
Age	-0.399	-0.400	-0.391	-0.399
	(0.056)	(0.056)	(0.055)	(0.052)
Age^2	0.004	0.004	0.004	0.004
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.125		
		(0.043)		
Income				-0.004
				(0.004)
Home value				-0.001
				(0.001)
R^2	0.0269	0.0270	0.0272	0.0274
Observations	167064	167064	167064	167064
Y-mean	9.602	9.602	9.602	9.602
	Panel B: No baseline gender difference in infection rate			
Member	3.260	3.341	3.326	3.294
	(1.940)	(1.927)	(1.926)	(1.958)
Female	0.335	0.318	0.312	0.314
	(0.322)	(0.322)	(0.323)	(0.330)
Member×Female	2.299	2.244	2.263	2.251
	(1.151)	(1.145)	(1.146)	(1.143)
Age	-0.418	-0.409	-0.395	-0.412
	(0.106)	(0.107)	(0.104)	(0.107)
Age^2	0.005	0.004	0.004	0.004
	(0.001)	(0.001)	(0.001)	(0.001)
Household size		-0.347		
		(0.103)		
Income				-0.006
				(0.005)
Home value				-0.003
				(0.002)
R^2	0.0583	0.0587	0.0589	0.0591
Observations	14437	14437	14437	14437
Y-mean	9.836	9.836	9.836	9.836
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S21: Household Transmission of Cancer as a Placebo Test

	Cancer infection dummy			
	(1)	(2)	(3)	(4)
Member_{cancer}	-0.159	-0.158	-0.168	-0.168
	(0.269)	(0.269)	(0.269)	(0.269)
Female	0.412	0.411	0.409	0.409
	(0.040)	(0.040)	(0.040)	(0.041)
$\text{Member}_{cancer}\times\text{Female}$	-0.542	-0.539	-0.534	-0.535
	(0.376)	(0.376)	(0.372)	(0.373)
Age	0.042	0.043	0.047	0.047
	(0.014)	(0.014)	(0.015)	(0.014)
Age^2	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Household size		-0.029		
		(0.014)		
Income				-0.000
				(0.001)
Home value				0.000
				(0.000)
Observations	80478	80478	80478	80478
R-square	0.0141	0.0141	0.0142	0.0142
Y-mean	1.173	1.173	1.173	1.173
ZIP FEs	Y	Y	Y	Y
Household size FEs			Y	Y

Notes: Cancer infection dummy is multiplied by 100. Robust standard errors are clustered at the zipcode level and reported in parentheses.

Table S22: Infectious Disease ICD-10-CM List

Disease code	Disease name
A02.0	Dysentery, dysenteric (infectious), Salmonella
A03.0	Dysentery, dysenteric (infectious), Schmitz(-Stutzer), Shiga(-Kruse)
A03.1	Dysentery, dysenteric (infectious), Flexner
A03.2	Dysentery, dysenteric (infectious), Boyd's
A03.3	Dysentery, dysenteric (infectious), Sonne
A03.8	Dysentery, dysenteric (infectious), specified type NEC
A03.9	Dysentery, dysenteric (infectious), bacillary
A04.0	Enteritis (infective) enteropathogenic
A04.1	Enteritis (infective) enterotoxigenic
A04.2	Enteritis (infective) enteroinvasive
A04.3	Enteritis (infective) enterohemorrhagic
A04.4	Enteritis (infective) enteroaggregative
A04.5	Enteritis (infective) Campylobacter
A04.6	Enteritis (infective) Yersinia enterocolitica
A04.71	Enteritis (infective) Clostridium difficile recurrent
A04.72	Enteritis (infective) Clostridium difficile not specified as recurrent
A04.8	Enteritis (infective) Clostridium perfringens
A04.9	Enteritis (infective) bacteria NOS
A06.0	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) amebic
A06.1	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) amebic chronic
A06.4	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) (sporadic) (tropical) abscess, liver
A07.0	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) balantidial
A07.2	Gastroenteritis (acute) (chronic) due to Cryptosporidium
A07.3	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) coccidial
A07.8	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) Embadomonas
A07.9	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) protozoal
A08.0	Gastroenteritis (acute) (chronic) rotaviral
A08.11	Gastroenteritis (acute) (chronic) type Norwalk
A08.39	Gastroenteritis (acute) (chronic) infantile
A08.4	Gastroenteritis (acute) (chronic) NEC
A08.8	Arthritis, arthritic (acute) (chronic) (nonpyogenic) (subacute) due to enteritis, infectious NEC A09 specified organism NEC A08.8
A09	Arthritis, arthritic (acute) (chronic) (nonpyogenic) (subacute) due to enteritis, infectious NEC
B15.0	Icterus infectious with hepatic coma
B15.9	Icterus infectious
B17.9	Hepatitis acute infectious

B18.9	Hepatitis acute infectious chronic
B27.89	Hepatomegaly in mononucleosis infectious specified NEC
B37.82	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) candidal
B55.0	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) leishmanial
B65.1	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) schistosomal
B78.0	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) strongyloidiasis
B82.0	Dysentery, dysenteric (catarrhal) (diarrhea) (epidemic) (hemorrhagic) (infectious) metazoal
D69.0	Purpura infectious
D69.2	Purpura
H01.9	Inflammation, inflamed, inflammatory (with exudation) eyelid dermatosis (infectious)
I50.1	Edema, edematous (infectious) lung failure, left ventricle
J00	Cold with influenza, flu, or grippe
J04.2	Laryngotracheitis (acute) (Infectional) (infective) (viral)
J09.X1	Bird influenza virus with pneumonia
J09.X2	Bird influenza virus
J09.X3	Bird influenza virus with digestive manifestations
J09.X9	Bird influenza virus with encephalopathy
J10.00	Identified influenza virus NEC with pneumonia (unspecified type)
J10.08	Identified influenza virus NEC with pneumonia specified type NEC
J10.89	Identified influenza virus NEC with specified manifestation NEC
J10.1	Identified influenza virus NEC
J10.2	Identified influenza virus NEC with digestive manifestations
J10.81	Identified influenza virus NEC with encephalopathy
J10.82	Identified influenza virus NEC with myocarditis
J10.83	Identified influenza virus NEC with otitis media
J11.00	Swine flu with pneumonia
J11.1	Influenza (bronchial) (epidemic) (respiratory (upper)) (unidentified influenza virus)
J12.2	Parainfluenza virus
J12.9	Viral, virus (broncho) (interstitial) (lobar)
J14	Hemophilus influenzae (broncho) (lobar)
J18.1	Lobar (disseminated) (double) (interstitial)
J18.9	Pneumonia (acute) (double) (migratory) (purulent) (septic) (unresolved)
J20.1	Haemophilus influenzae
J20.4	Virus parainfluenzae
J20.9	Croup, croupous bronchial
J21.9	Bronchiolitis (acute) (infective) (subacute)
J31.1	Catarrh, catarrhal nasobronchial
J31.2	Catarrh, catarrhal throat
J32.9	Sinusitis (accessory) (chronic) (hyperplastic) (nasal) (nonpurulent) (purulent)
J40	Bronchitis (diffuse) (fibrinous) (hypostatic) (infective) (membranous)

J68.1	Edema, edematous (infectious) lung chemical
J70.0	Edema, edematous (infectious) due to radiation
J70.1	Edema, edematous (infectious) lung due to radiation
J70.8	Edema, edematous (infectious) lung due to external agent
J70.9	Edema, edematous (infectious) lung due to specified NEC
J81.0	Edema, edematous (infectious) lung acute
J81.1	Edema, edematous (infectious) lung
J98.9	Disease, diseased respiratory (tract)
K52.0	Gastroenteritis (acute) (chronic) due to radiation
K52.1	Gastroenteritis (acute) (chronic) due to drugs
K52.21	Gastroenteritis (acute) (chronic) due to food protein-induced enterocolitis syndrome
K52.22	Gastroenteritis (acute) (chronic) due to food protein-induced enteropathy
K52.29	Gastroenteritis (acute) (chronic) due to dietetic
K52.81	Gastroenteritis (acute) (chronic) due to eosinophilic
K52.89	Gastroenteritis (acute) (chronic) due to specified NEC
K52.9	Gastroenteritis (acute) (chronic)
K75.9	Hepatitis
L30.3	Infectious eczematoid
L30.9	Dermatitis (eczematous)
L98.9	Dermatosis
P23.6	Congenital (infective) P23.9 due to Haemophilus influenzae
P23.9	Congenital (infective)
R05	Cough (affected) (chronic) (epidemic) (nervous)
R19.7	Diarrhea, diarrheal (disease) (infantile) (inflammatory)
T75.1	Edema, edematous (infectious) lung due to near drowning
T75.89	Exposure (to) intestinal infectious disease
Z11.0	Screening (for) intestinal infectious
Z20.01	Contact (with) intestinal infectious disease NEC
Z20.09	Contact (with) intestinal infectious disease NEC Escherichia coli (E. coli)
Z22.0	Carrier (suspected) of bacterial disease NEC intestinal infectious NEC typhoid
Z22.1	Carrier (suspected) of bacterial disease NEC intestinal infectious NEC
Z86.13	Personal (of) disease or disorder (of) infectious malaria

Table S23: Top 10 Health Issues by Gender

(a) Top 10 in Women

Order	Infection	Percentage
1	Cough, Cold, Sore Throat	24.81
2	Respiratory Infections	8.83
3	Hypertension	3.23
4	Headache or Cephalalgia	2.20
5	Diabetes	2.16
6	Physical Injury by Accident	2.15
7	Diarrhea	1.45
8	Gastritis or Gastric Ulcer	1.21
9	Colitis	0.93
10	Allergies	0.78

(b) Top 10 in Men

Order	Infection	Percentage
1	Cough, Cold, Sore Throat	21.07
2	Respiratory Infections	7.97
3	Physical Injury by Accident	3.31
4	Diarrhea	1.48
5	Diabetes	1.35
6	Hypertension	1.13
7	Headache or Cephalalgia	0.94
8	Gastritis or Gastric Ulcer	0.93
9	Allergies	0.64
10	Asthma	0.63

Notes: Data comes from the National Health and Nutrition Survey in Mexico (ENSANUT) round 2012. Percentage shows the share of health problems that happened within the 2 weeks of the interview excluding "others".

Table S24: Top 10 Health Issues with Medical Visit by Gender

(a) Top 10 in Women

Order	Infection	Percentage
1	Cough, Cold, Sore Throat	13.41
2	Respiratory Infections	5.28
3	Hypertension	2.59
4	Diabetes	1.80
5	Physical Injury by Accident	1.44
6	Headache or Cephalalgia	0.95
7	Gastritis or Gastric Ulcer	0.92
8	Diarrhea	0.87
9	Colitis	0.74
10	Asthma	0.64

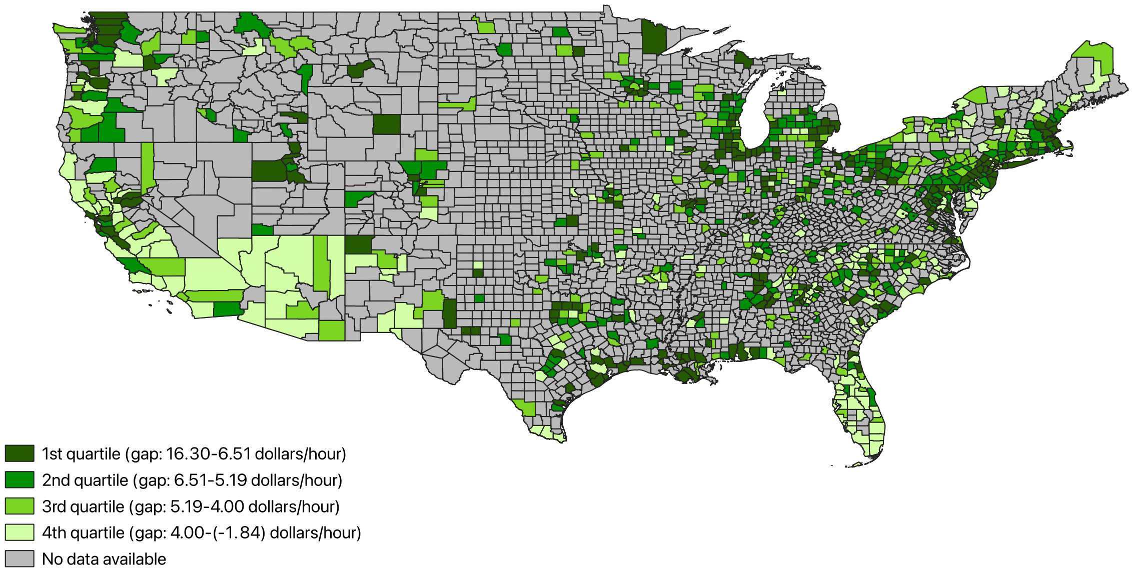
(b) Top 10 in Men

Order	Infection	Percentage
1	Cough, Cold, Sore Throat	10.88
2	Respiratory Infections	4.60
3	Physical Injury by Accident	2.18
4	Diabetes	1.15
5	Hypertension	0.88
6	Diarrhea	0.84
7	Gastritis or Gastric Ulcer	0.58
8	Asthma	0.51
9	Allergies	0.43
10	Dental Disease	0.38

Notes: Data comes from the National Health and Nutrition Survey in Mexico (ENSANUT) round 2012. Percentage shows the share of health problems that happened within the 2 weeks of the interview and were consulted by a medical professional (excluding "others").

S6 Additional Figures

Figure S1: Distribution of gender gap in wage and home production time



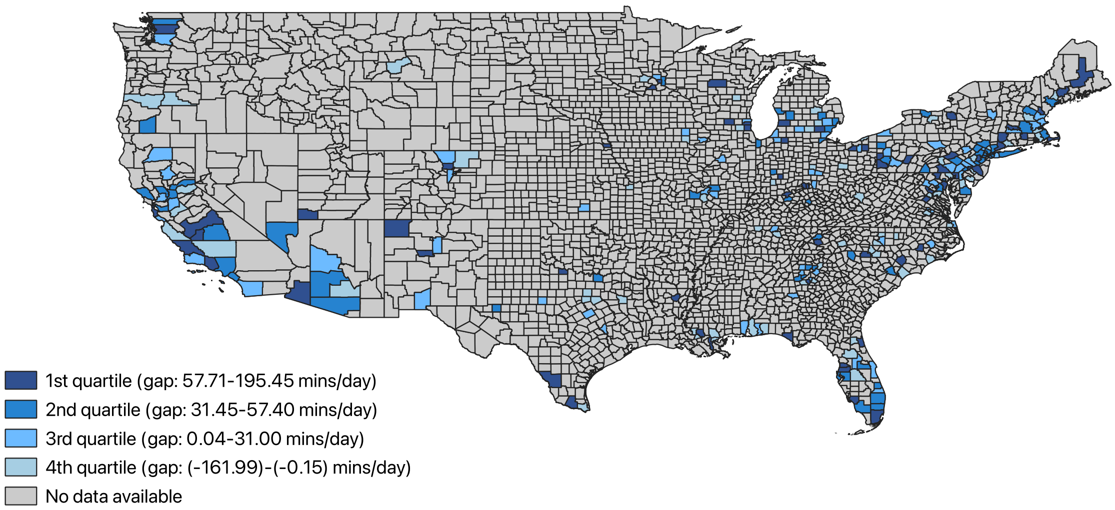
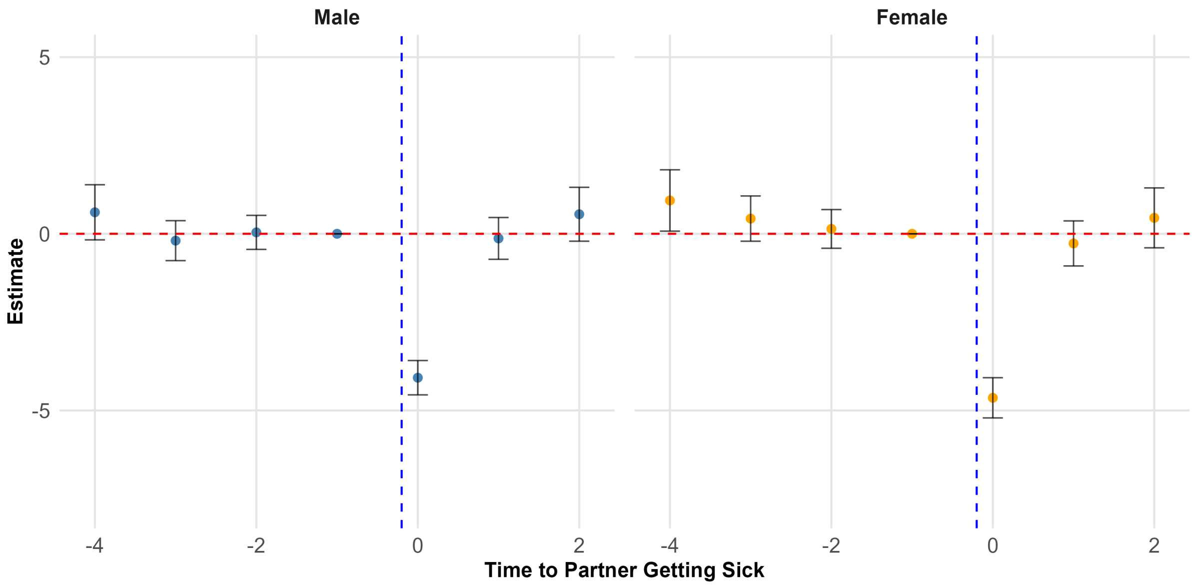
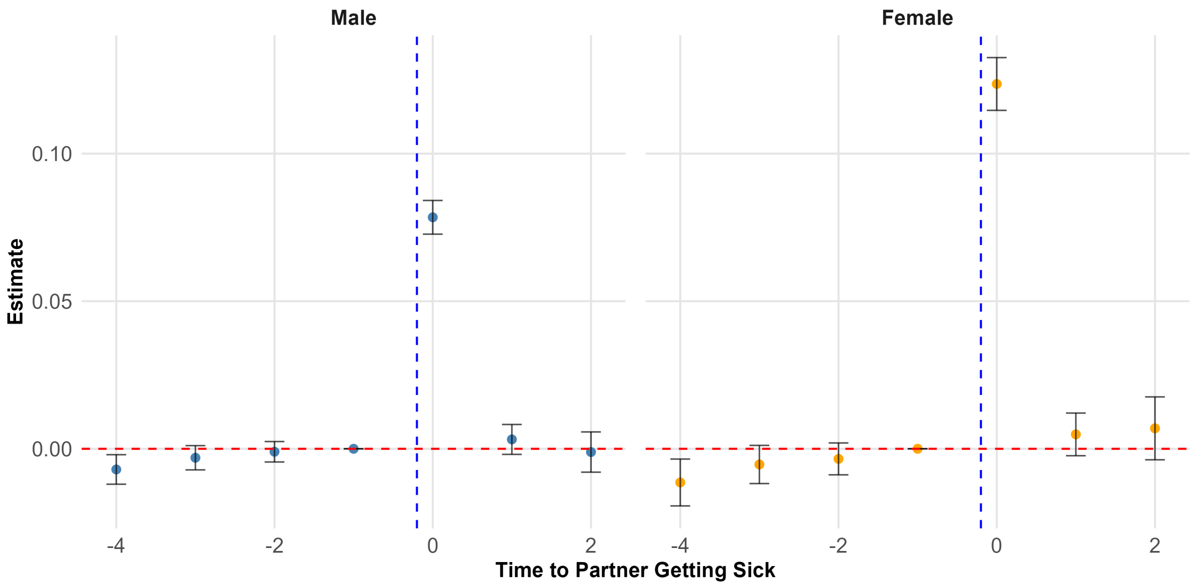


Figure S2: Event Study: Working Hours



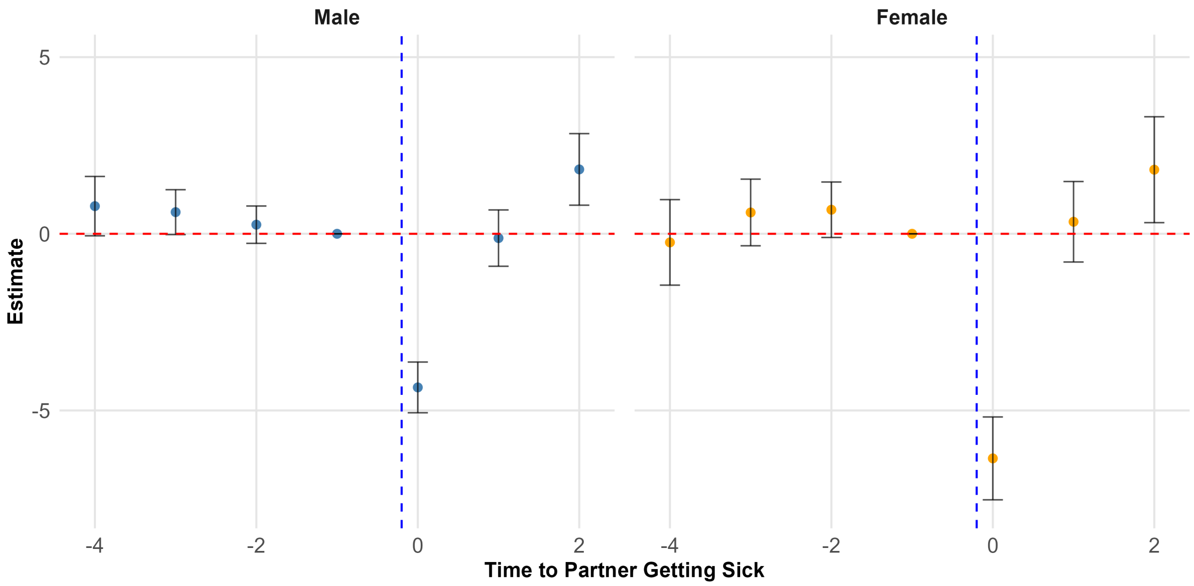
Note: The figure shows event study estimates of the impact of a partner's illness on working hours, separated by gender. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue precedes the interview time with the event

Figure S3: Event Study: Sickness



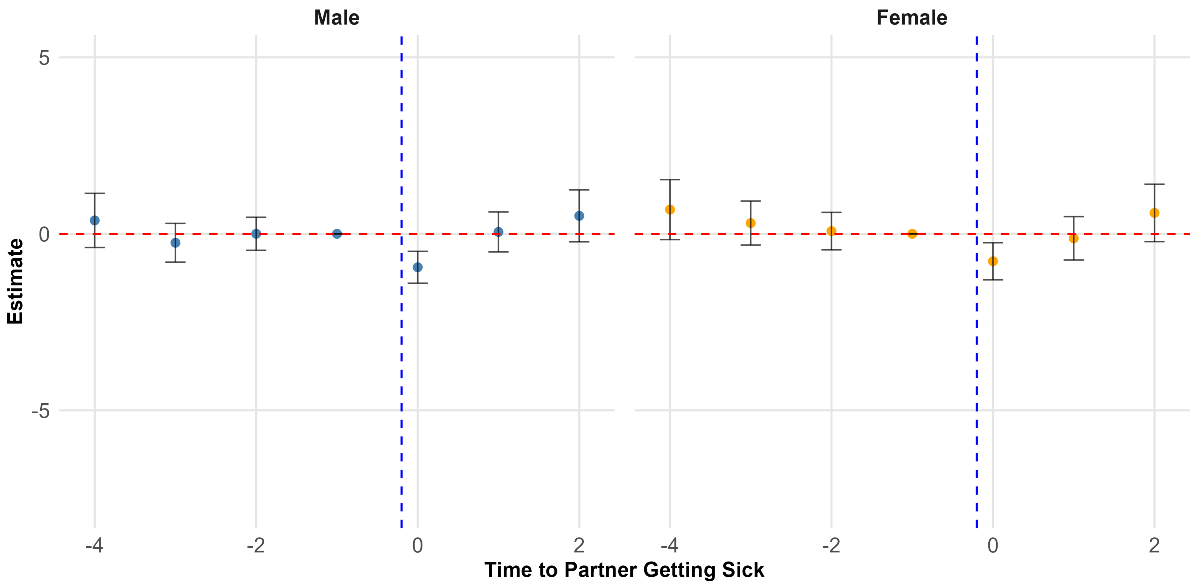
Note: The figure shows event study estimates of the impact of a partner's illness on probability of own sickness, separated by gender. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue precedes the interview time with the event

Figure S4: Event Study: Working Hours Conditional on Full Time Employment



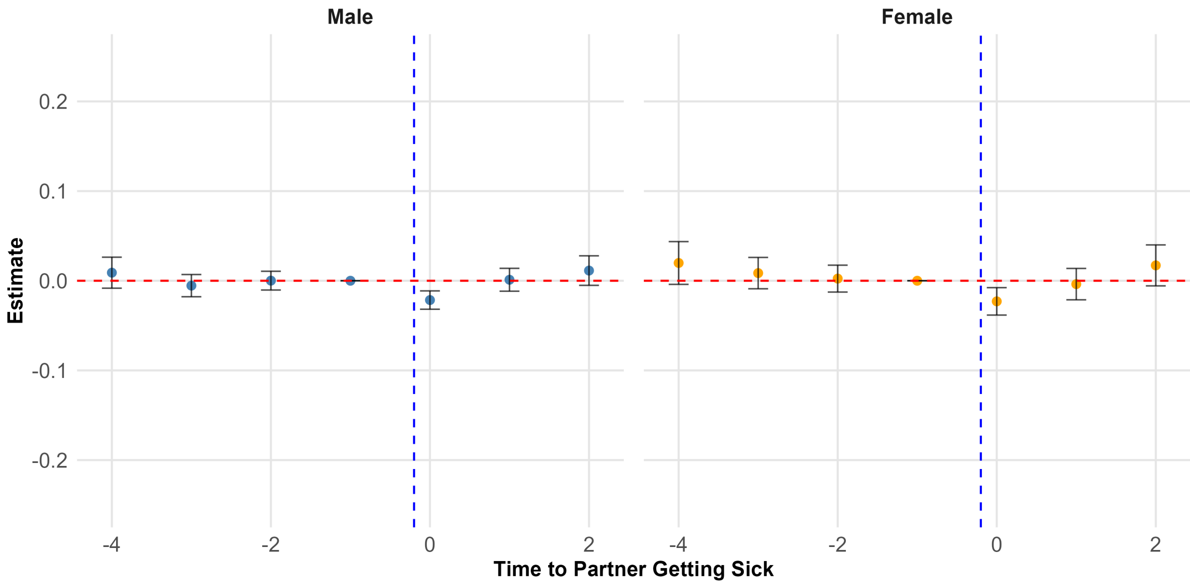
Note: The figure shows event study estimates of the impact of a partner's illness on working hours, separated by gender, using only sample of individuals who have had full job pre-event. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue precedes the interview time with the event

Figure S5: Event Study: Working Hours Conditional on Not Getting Sick



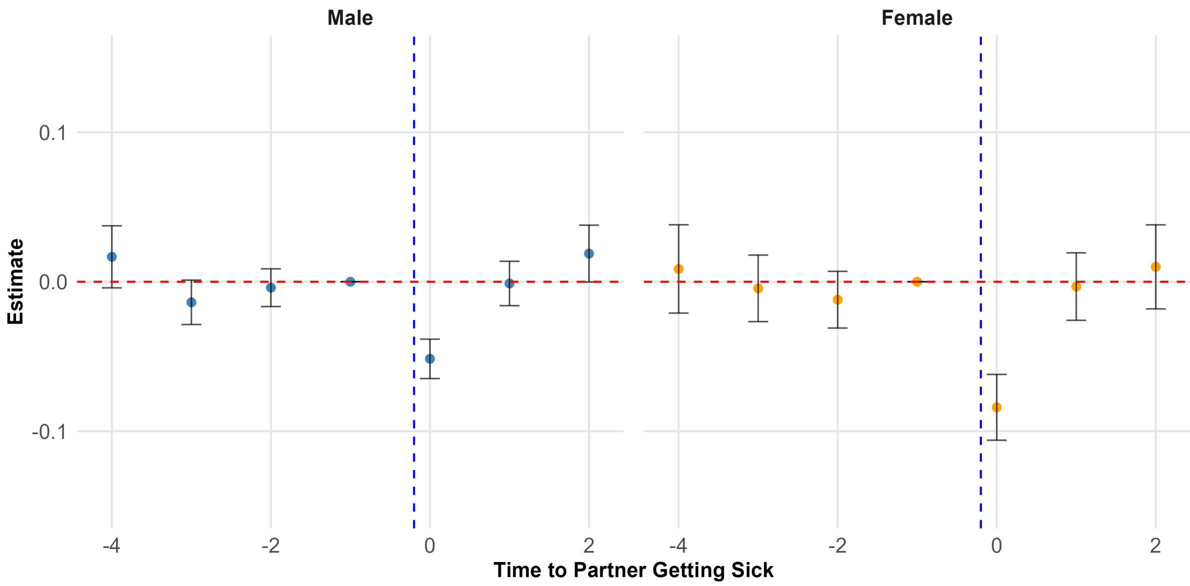
Note: The figure shows event study estimates of the impact of a partner's illness on working hours, separated by gender, using only sample of individuals have not been sick themselves. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue precedes the interview time with the event

Figure S6: Poisson Event Study: Working Hours Conditional on Not Getting Sick

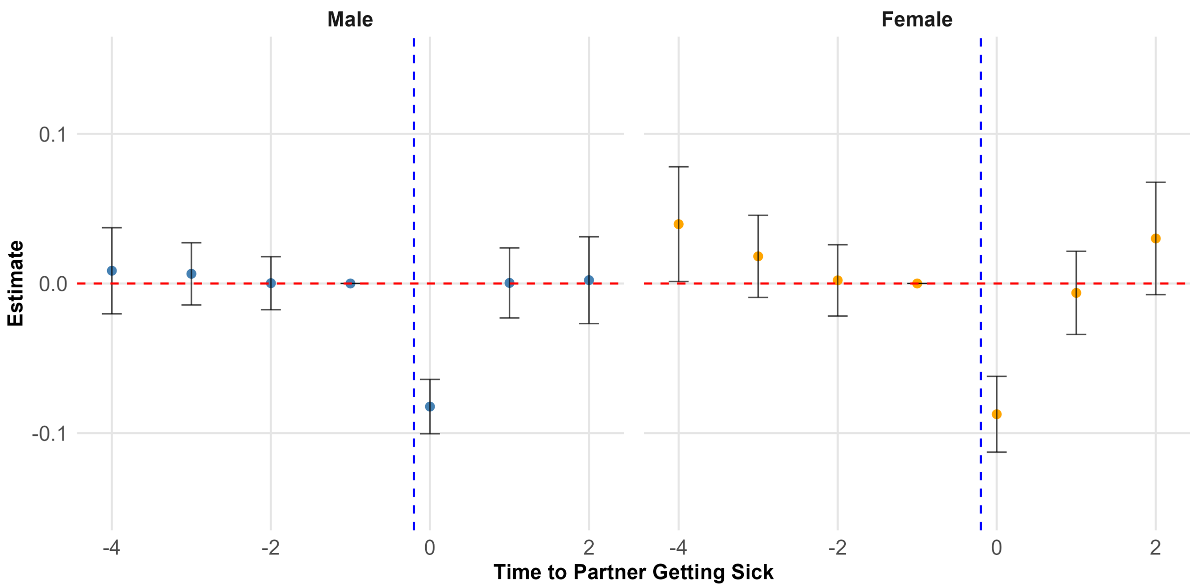


Note: The figure shows event study estimates of the impact of a partner's illness on working hours, separated by gender, using only sample of individuals have not been sick themselves and a Poisson regression. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue precedes the interview time with the event

Figure S7: Poisson Event Study: Working Hours and Informality



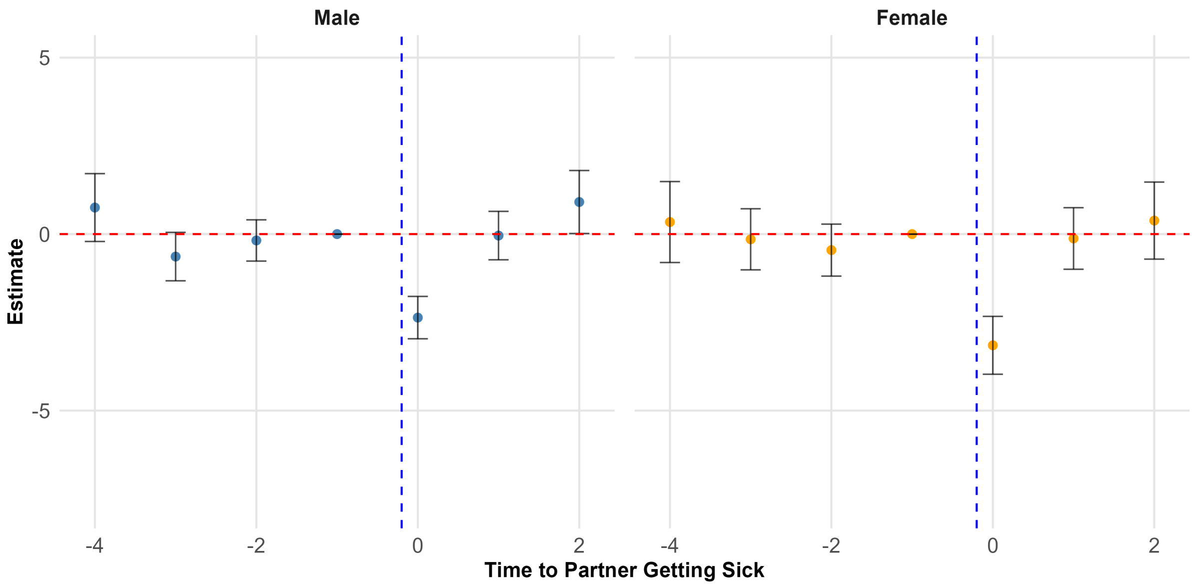
(a) Formal Sector



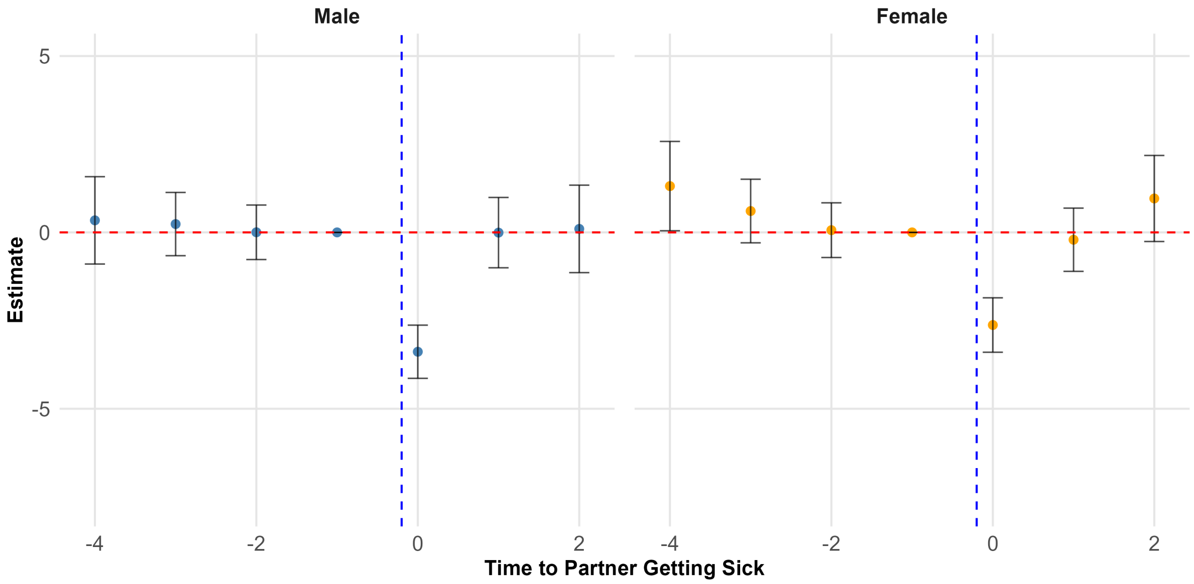
(b) Informal Sector

Note: The figure shows event study estimates of the impact of a partner's illness on working hours using a Poisson model, separated by gender and whether they work in the formal or informal sector. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue line precedes the interview time with the event.

Figure S8: Event Study: Working Hours and Informality



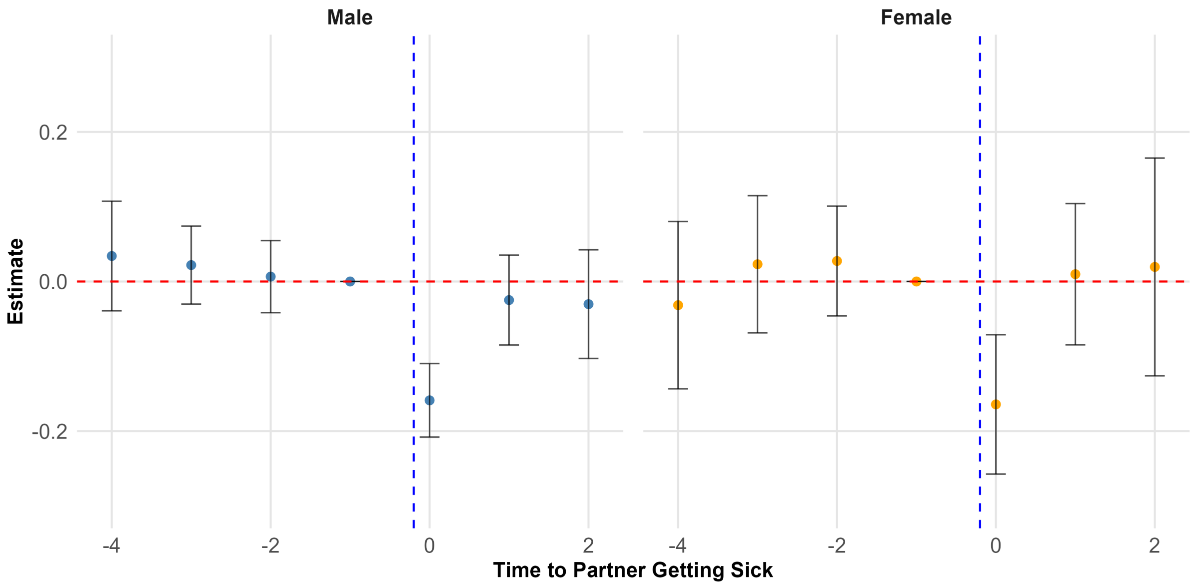
(a) Formal Sector



(b) Informal Sector

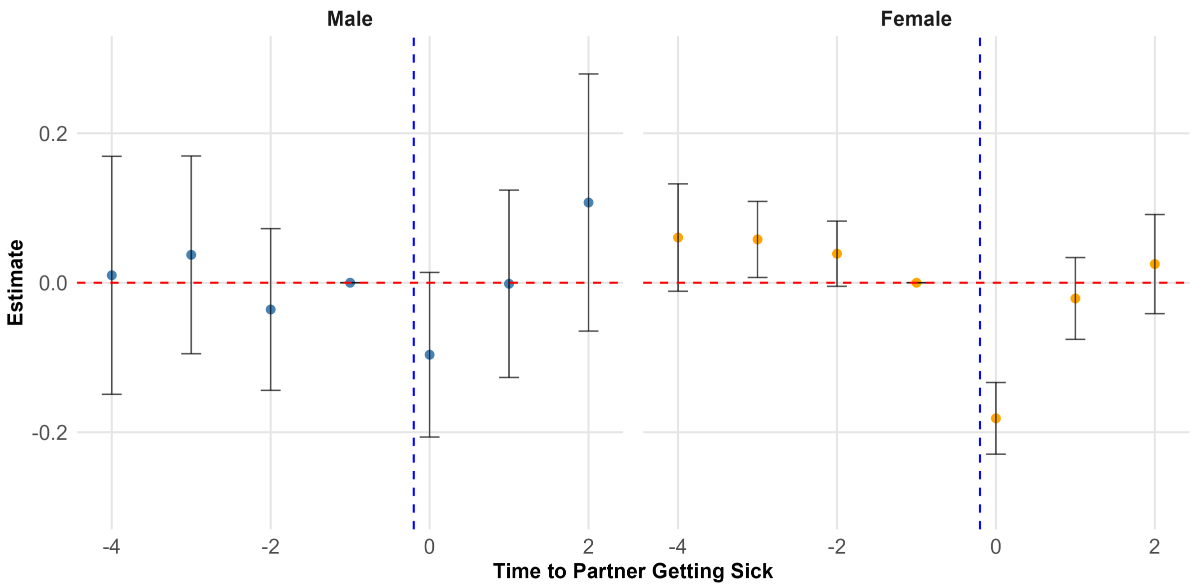
Note: The figure shows event study estimates of the impact of a partner's illness on working hours, separated by gender and whether they work in the formal or informal sector. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue line precedes the interview time with the event.

Figure S9: Poisson Event Study: Working Hours and Elderly Female Presence



Note: The figure shows event study estimates of the impact of a partner's illness on working hours, separated by gender and only for individuals who live with an elderly female. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue precedes the interview time with the event

Figure S10: Poisson Event Study: Working Hours and Elderly Male Presence



Note: The figure shows event study estimates of the impact of a partner's illness on working hours, separated by gender and only for individuals who live with an elderly male. The x-axis indicates time (quarters) relative to the event, and the y-axis shows changes in working hours. Dots represent point estimates, with 95% confidence intervals, and the dashed blue precedes the interview time with the event

Figure S11: Increased Infection Rate for Men and Women (top) and its Gender Gap (bottom)