



Universidade Federal de Pernambuco
Programa de Pós-Graduação em Economia - PIMES

Manoel Tiago Lima da Silva

Efeito das firmas sobre o contágio da COVID-19 em Recife

Recife

2023

Manoel Tiago Lima da Silva

Efeito das firmas sobre o contágio da COVID-19 em Recife

Dissertação submetida ao Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco (área de concentração: teoria econômica), como requisito parcial para a conclusão do Mestrado em Economia.

Orientadora: Prof^ª. Dra. Tatiane Almeida de Menezes

Recife

2023

Catálogo na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

S586e Silva, Manoel Tiago Lima da
Efeito das firmas sobre o contágio da COVID-19 em Recife / Manoel
Tiago Lima da Silva. - 2023.
38 folhas: il. 30 cm.

Orientadora: Prof.^a Dra. Tatiane Almeida de Menezes.
Dissertação (Mestrado em Economia) – Universidade Federal de
Pernambuco, CCSA, 2023.
Inclui referências.

1. Covid-19. 2. Ambiente de trabalho. 3. Modelos de probabilidade. I.
Menezes, Tatiane Almeida de (Orientadora). II. Título.

336 CDD (22. ed.) UFPE (CSA 2023 – 028)

MANOEL TIAGO LIMA DA SILVA

EFEITO DAS FIRMAS SOBRE O CONTÁGIO DA COVID-19 EM RECIFE

Dissertação ou Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco (área de concentração: teoria econômica), como requisito parcial para a obtenção do título de Mestre em Economia.

Aprovado em: 15/03/2023.

BANCA EXAMINADORA

Prof^o. Dr. Tatiane Almeida de Menezes (Orientador)
Universidade Federal de Pernambuco

Prof^o. Dr. Diego Firmino Costa da Silva (Examinador Interno)
Universidade Federal de Pernambuco

Prof^o. Dr. Julyan Gleyvison Machado Gouveia Lins (Examinador Externo)
Universidade Federal da Bahia

Resumo

Este trabalho tem como objetivo verificar qual é o efeito das firmas sobre o contágio do indivíduo pela COVID-19. Para isso, será estimado um modelo de probabilidade (*probit*) que terá como variável dependente uma *dummy* que indica se o indivíduo foi ou não infectado pelo vírus. A variável de interesse será a proporção de indivíduos na firma que já foram diagnosticados com a COVID-19. Para tratar uma possível endogeneidade presente no modelo, optou-se por adotar a estratégia empírica descrita como variável instrumental (IV). O instrumento utilizado será o número de pessoas que testaram positivo no bairro onde os colegas de trabalho moram, desde que estes colegas não residam no mesmo bairro que o trabalhador em questão. Os resultados mostram que o ambiente de trabalho possui um alto potencial de transmissão, pois o acréscimo de 1 ponto percentual no número de infectados das firmas, implicará em um incremento de cerca 2,2% na probabilidade de contágio do indivíduo.

Palavras-chave: COVID-19; Ambiente de Trabalho; Modelo de Probabilidade.

Abstract

This work aims to verify the effect of firms on the individual's contagion by COVID-19. For this, a probability model (probit) will be estimated, which will have as a dependent variable a dummy that indicates whether or not the individual was infected by the virus. The variable of interest will be the proportion of individuals in the firm who have already been diagnosed with COVID-19. To address a possible endogeneity present in the model, it was decided to adopt as an empirical strategy described as an instrumental variable (IV). The instrument used will be the number of people who tested positive in the neighborhood where co-workers live, as long as these co-workers do not reside in the same neighborhood as the worker in question. The results show that the work environment has a high transmission potential, since the increase of 1 percentage point in the number of infected companies will imply an increase of about 2.2% in the individual's probability of contagion.

Keywords: COVID-19; Workplace; Probability Model.

Sumário

1. INTRODUÇÃO	7
2. REVISÃO DE LITERATURA	9
2.1 COVID-19 e fatores socioeconômicos	9
2.2 COVID-19 e dimensão ocupacional	11
3. METODOLOGIA	16
3.1 Dados	16
3.3 Estratégia Empírica	17
4. RESULTADOS	21
4.1 Análise descritiva	21
4.2 Resultados dos modelos e Discussão	25
5. CONSIDERAÇÕES FINAIS	33
REFERÊNCIAS	34

1. INTRODUÇÃO

No contexto da infecção por COVID-19, o local de trabalho dos indivíduos foi considerado um dos principais veículos de transmissão para o vírus, pois o contágio ocorre de uma pessoa infectada para outra. Conforme relatório da Organização Mundial da Saúde (OMS) e Organização Internacional do Trabalho (OIT), alguns fatores são relevantes para a disseminação nesse meio, tais como: i) fatores ocupacionais da firma: como a proximidade física, ventilação inadequada, possibilidade de trabalho remoto, dentre outros; ii) fatores sociodemográficos: como a idade, sexo, raça, renda e status de imigrante; iii) fatores paraocupacionais: como viagem ao trabalho, encontros sociais entre colegas fora do trabalho; iv) intervenções para diminuição contágio. Esse entendimento de como o vírus se propaga no ambiente de trabalho pode contribuir para que as políticas sejam aprimoradas (OMS e OIT, 2021).

Ainda conforme este relatório o Sars-Cov-2 se transmite principalmente entre indivíduos que estão a curta distância um do outro. Essa transmissão pode ocorrer em pequenas partículas lançadas quando as pessoas tosse, espirram, falam, cantam ou respiram. Outra forma de contrair o vírus é quando alguém inala partículas infectadas suspensas no ar ou quando alguém as coloca em contato com olhos, nariz ou boca. O vírus também pode se espalhar mais facilmente em locais pouco ventilados ou lotados de pessoas (OMS e OIT, 2021).

Em um primeiro momento, devido à ausência de tratamentos farmacológicos, autoridades sanitárias de vários países recomendaram a utilização de medidas como distanciamento social, uso de máscaras e medidas mais severas como o *lockdown*. Para que seja possível a elaboração de estratégias eficazes de redução do contágio, é necessário entender como essa doença se transmite nos distintos locais onde as pessoas estão inseridas. Isso porque cada ambiente possui características próprias que podem torná-lo mais ou menos propício para o contágio.

Sendo assim, é relevante verificar se trabalhadores infectados de uma mesma firma, potencializam o contágio pelo Sars-COV-2, pois se trata de um local compartilhado no qual os indivíduos estão expostos em boa parte do dia. Para isso, neste trabalho será estimado um modelo de probabilidade que visa captar o impacto que as firmas causam na contaminação do indivíduo. Será considerada como variável dependente, se um indivíduo foi ou não infectado pelo vírus. A variável de interesse será a proporção de indivíduos na firma que já foram diagnosticados com a COVID-19.

A relação entre a proporção de infectados na firma e a probabilidade de um funcionário sofrer o contágio está positivamente correlacionada, sendo determinada simultaneamente. A presença de endogeneidade na relação entre essas variáveis ocasiona a estimação de parâmetros viesados para o modelo. Desta forma, para tentar identificar o impacto da prevalência da COVID-19 na firma sobre a chance de alguém que nela trabalha ter a referida doença, optou-se por adotar como estratégia empírica o método descrito como variável instrumental (IV). O instrumento utilizado será o número de pessoas que testaram positivo no bairro onde os colegas de trabalho moram, desde que estes colegas não residam no mesmo bairro que o trabalhador em questão, este instrumento segue uma abordagem semelhante às propostas por Kelejian e Robinson (2003), Reed (2015) e em um trabalho aplicado por Feng e Humphreys (2016).

Os dados relativos à saúde foram disponibilizados pela Secretaria Estadual de Saúde de Pernambuco, num banco aqui denominado de SUS_COVID. Por sua vez, os dados socioeconômicos e ocupacionais foram obtidos no banco de dados da Relação Anual de Informações Sociais (RAIS). Desse modo, devido a disponibilidade dos dados, o estudo será restrito aos trabalhadores do mercado formal da cidade do Recife. Ambas bases de dados permitirão a realização de um estudo baseado em um conjunto de características socioeconômicas a nível de indivíduos. Até o momento, não foram encontrados estudos sobre a COVID-19 que trazem esse nível de desagregação nos dados, o que permite uma contribuição inédita para a literatura. Além disso, as observações da base de dados são georreferenciadas, possibilitando uma abordagem que contemple a dimensão espacial, que é de fundamental importância quando trabalhamos com contágio por vírus.

De modo geral, os resultados encontrados apontam que firmas que possuem mais trabalhadores contaminados, tornam o indivíduo mais propenso a também ser contaminado. Além disso, características socioeconômicas como raça, gênero e renda também são importantes para explicar a probabilidade de contágio.

Além desta introdução, esse trabalho contém mais 3 seções, na próxima será feita uma breve revisão de literatura com algumas referências nacionais e internacionais sobre o tema. Na sequência, será definida a metodologia com os dados e estratégia empírica a serem utilizadas no trabalho. Em seguida, serão apresentados os resultados e discussão e por fim as considerações finais.

2. REVISÃO DE LITERATURA

2.1 COVID-19 e fatores socioeconômicos

O estudo de fatores de risco que podem afetar a saúde dos indivíduos vem sendo desenvolvido por estudiosos de diversas áreas como epidemiologia, economia, sociologia, dentre outras. Esses pesquisadores buscam entender como estruturas sociais, instituições, fatores econômicos e relacionamentos entre pessoas influenciam a saúde da população e dos indivíduos.

Nesse sentido, o status socioeconômico aparece como um importante fator explicativo, pois abrange múltiplas dimensões da vida do indivíduo como educação, emprego e renda. Um baixo nível de status socioeconômico está associado com a maioria dos casos de morte prematura proveniente de doenças infecciosas (BERKMAN, KAWACH e GLYMOUR, 2014).

De acordo com Berkman, Kawachi, Glymour (2014) uma das abordagens teóricas para explicar a relação entre condições socioeconômicas e saúde foi proposta por Link e Phelan (1995). Conforme essa teoria, indivíduos com alto nível de status socioeconômico possuem uma grande variedade de recursos tais como dinheiro, conhecimento, prestígio, poder e conexões que ajudam evitar a exposição a fatores de risco na sua saúde. Link e Phelan (1995) também argumentam que é importante observar não apenas fatores de risco individuais, devendo-se atentar também ao motivo pelo qual em algumas populações as pessoas estão mais propensas a ficarem expostas a esses fatores do que em outras.

Pires, Carvalho e Xavier (2020) avaliam a relação entre a pobreza da população e a existência de uma série de fatores que a torna ainda mais vulnerável ao contágio em pandemias como a da COVID-19. Dentre esses fatores estão: o uso de transporte público, maior número de moradores por domicílio, saneamento básico e o acesso à saúde inadequados e a dificuldade de manter o isolamento social sem grande perda de renda ou do emprego.

Desse modo, a circulação das pessoas e contaminação pelo vírus não ocorre de maneira homogênea nos locais afetados, estando condicionadas às características econômicas e sociais mencionadas acima. Nessa linha, o estudo sobre a pandemia e seus efeitos sobre a saúde da população deve levar em consideração a forte desigualdade presente na sociedade.

A Pesquisa Nacional de Saúde de 2019 nos ajuda a ter uma noção da desigualdade no acesso à saúde existente no Brasil. Apenas 16,1% da população sem instrução ou com

ensino fundamental incompleto tem acesso a algum plano de saúde, enquanto entre os que possuem diploma de nível superior o acesso é de 67,7%. Entre os que ganham de 1 a 2 salários mínimos, existe uma cobertura de 32,4% de plano de saúde médico; por outro lado, entre os que ganham mais de 5 salários mínimos a cobertura é de 86,8%.

Marmot (2005) considera que para reduzir a desigualdade na saúde entre países é importante, além do fortalecimento dos sistemas de saúde e redução da pobreza, focar nos determinantes sociais da saúde. Para isso, é importante que as políticas públicas sejam implementadas de forma global e não apenas por entidades responsáveis pela saúde. Dessa forma, apenas identificar e suprir as necessidades materiais da população não é suficiente, deve-se considerar as circunstâncias nas quais as pessoas vivem e trabalham.

Consolazio et al. (2021) buscaram verificar, em um estudo realizado para a região da Lombardia na Itália, se características socioeconômicas da vizinhança dos indivíduos afetaram o risco de contágio pelo coronavírus. Foi utilizado um modelo de regressão logística multinível que considera características individuais e de área. Os resultados mostraram que há um efeito do contexto socioeconômico sobre o risco de contágio, mesmo controlando por variáveis individuais.

Utilizando um modelo de regressão logística, Quan et al. (2021) avaliaram a associação entre condições sociodemográficas dos pacientes e condições precárias da vizinhança sobre a frequência de morte, ventilação mecânica invasiva e admissão em unidades de tratamento intensivo. Os resultados mostram que pacientes de vizinhanças que possuem menor renda estão mais propensos a serem admitidos em UTIs e serem submetidos a ventilação mecânica invasiva.

Mena et al. (2021) investigam a influência de fatores socioeconômicos no impulsionamento da síndrome aguda respiratória (Sars-Cov-2). A taxa de infecções mortais foi maior em municípios de baixa renda devido a comorbidades e falta de acesso ao sistema de saúde. Além disso, os resultados mostram que as desigualdades socioeconômicas afetam os mais jovens desproporcionalmente.

Rutter et al. (2012) analisa a variação da mortalidade na mortalidade entre grupos socioeconômicos na Inglaterra devido à pandemia do vírus da influenza (H1N1). A condição socioeconômica dos indivíduos foi determinada a partir das localizações por meio do código postal. Os resultados mostraram que os grupos de pessoas que vivem em regiões socioeconômicas mais precárias e pessoas que vivem em regiões urbanas apresentaram uma taxa de mortalidade maior que os demais grupos

Muitos autores chamam a atenção para ausência de política que deem enfoque a situação socioeconômica dos indivíduos. Nesse contexto, conforme Procópio (2020) torna-se importante a adoção de estudos que considerem essa dimensão, “como é o caso do auxílio emergencial, e de garantia da segurança alimentar e nutricional por meio de acesso a uma alimentação saudável para minimizar a expansão dos casos de COVID-19”.

2.2 COVID-19 e dimensão ocupacional

Os impactos das medidas adotadas para redução da transmissão da COVID-19 atingiram de forma diferenciada os distintos setores da economia. Esses impactos foram sentidos em todos os setores, mas há um destaque para os trabalhadores autônomos e informais. Enquanto no setor privado formal cerca de 17% dos trabalhadores foram para o desemprego ou inatividade, no setor privado informal esse número foi de 35% entre o primeiro trimestre e Maio de 2020. Houve também perda de renda, no setor privado informal mais de 30% receberam menos da metade da renda habitual, enquanto no setor formal 7,8% tiveram essa perda (Carvalho e Nogueira, 2020). Por outro lado, o setor de administração pública e agricultura foram os menos afetados como mostra o trabalho de Corseuil Ramos e Russo (2020).

Esse quadro reforça a necessidade de políticas que visem diminuir a perda de emprego, renda e a vulnerabilidade dos trabalhadores. Para a conservação do nível de emprego, por exemplo, foi proposto por Hecksher e Corseuil (2020) a desoneração de jornadas de trabalho mais curtas como forma de evitar demissões diante de uma demanda menor por horas de trabalho. Já o auxílio emergencial, programa de renda mínima do governo federal, permitiu que a renda das famílias se mantivesse num patamar mínimo de subsistência.

Em reforço a proteção ao trabalho, o Supremo Tribunal Federal (STF) suspendeu a eficácia de dois dispositivos da Medida Provisória (MP) 927/2020: o artigo 29 que exigia nexo causal para comprovação da COVID-19 como doença ocupacional e o artigo 31 que limitava a atuação de auditores fiscais apenas a orientação das atividades relacionadas ao trabalho. Essas decisões foram importantes no contexto da pandemia, pois evitaram que medidas arbitrárias por parte dos empregadores expusessem ainda mais trabalhadores em situação de fragilidade socioeconômica.

No trabalho de Costa Barbosa e Hecksher (2021) vemos que os trabalhadores são afetados de forma bastante desigual, essa desigualdade se dá em várias dimensões como

sexo, raça, idade, ocupação etc. Sendo que os mais afetados são aqueles que já se encontram em situações mais vulneráveis intensificando ainda mais o problema. Dentre os negros, por exemplo, aumentou a transição para inatividade e houve uma redução maior da ocupação se comparados com os brancos. Em relação a idade em 2015 os adultos tinham ocupação de 7,3 pontos percentuais acima dos jovens, em 2020 essa diferença se eleva para 12,3 p.p. As mulheres, por sua vez, também foram mais afetadas do que os homens. Souza (2021) traz evidências de que a massa salarial dos negros caiu mais do que a dos brancos entre o primeiro e o segundo trimestre de 2020 e também houve redução maior nos postos de trabalho.

Silva e Vaz (2020) apontam que houve também um aumento do número de jovens desalentados, que são aqueles que abrem mão de procurar emprego por não terem expectativa encontrar. Além desse impacto no mercado de trabalho, a própria reinserção é dificultada na medida em que a construção de habilidades é prejudicada, pois não havia meios adequados de prosseguir com educação e formação profissional, fundamentais para ampliar as chances de conquistar um emprego. Os autores ainda consideram que houve um agravamento da situação dos jovens que não estudam e não trabalham, pois muitos interromperam os estudos e pararam de buscar emprego.

No campo da educação, conforme Paes de Barros et al (2021) houve perdas de longo prazo na aprendizagem dos estudantes. Este estudo conclui que os estudantes que concluíram a 2ª série do ensino médio em 2020, iniciaram a 3ª série com uma proficiência de cerca de 10 pontos menor no SAEB (Sistema de Avaliação da Educação Básica) em português e matemática do que teriam num cenário sem pandemia. Visando reduzir estas perdas, os autores sugerem que algumas medidas poderiam ser adotadas como por exemplo: promover o maior engajamento dos estudantes com o ensino remoto, adotar rapidamente alguma forma de ensino híbrido.

As características associadas à ocupação e ambiente de trabalho também vêm sendo descritas na literatura como fatores importantes para explicação do risco de infecção dos indivíduos. A transmissão do SARS-CoV-2 ocorre em uma multiplicidade de locais de trabalho, sendo que alguns desses setores possuem características particulares que o tornam mais ou menos propenso para essa transmissão. Nesse sentido, muitos estudos apontam o ambiente de trabalho como um fator relevante para a contaminação dos indivíduos pelo vírus da Covid-19.

Segundo Carlsten et. al (2021) desde a pandemia de 1918-1919, que matou aproximadamente 75 milhões de pessoas em todo mundo, já havia um gradiente

socioeconômico para as mortes, e os trabalhadores de setores de maior aglomeração estavam expostos a maiores riscos. Na pandemia da COVID-19, evidências mostram que trabalhadores de setores essenciais apresentam maiores riscos de contrair o vírus Sars-Cov-2.

As disparidades no contágio e morte por COVID-19 estão relacionadas a inúmeros fatores, incluindo a natureza e o perigo dos trabalhos, as condições prévias de saúde e fatores socioeconômicos. As implicações da Covid-19 para a força de trabalho como um todo são complexas e multifacetadas, pois as características individuais de suscetibilidade como sexo, idade avançada e comorbidades, interagem com os fatores de risco presentes no ambiente de trabalho (CARLSTEN et. al, 2021).

Em uma revisão de literatura sobre os impactos da COVID-19 sobre a saúde do trabalhador Aguiar, Sarquis e Miranda (2021) identificaram que entre as possíveis sequelas, apontam-se as neurológicas, respiratórias, psicológicas, vasculares, gastrointestinais, cutâneas, e até as relacionadas ao olfato e paladar. Os autores consideram que estas sequelas podem estar relacionadas ao trabalho, interferem na piora da qualidade de Vida dificultando as atividades diárias e diminuindo a produtividade dos trabalhadores.

Nesse sentido, Mutambudzi et al. (2021) investigaram, através de uma regressão Poisson, o risco de casos severos de COVID-19 por grupos ocupacionais no Reino Unido. Os resultados mostram que trabalhadores de setores essenciais têm maiores riscos de apresentarem este quadro mais grave. Dentre esses trabalhadores o subgrupo de não-brancos apresentou risco ainda maior de contrair a forma grave da doença. Esse estudo também mostra que, conforme vem sendo amplamente constatado na literatura, trabalhadores da área da saúde são os mais expostos.

Leso, Fontana e Iavicoli (2021) efetuaram uma pesquisa nos repositórios *PubMed*, *Scopus*, e *ISI Web of Science* até 20 de dezembro de 2020, em busca da literatura a respeito de condições de suscetibilidade e vulnerabilidade associadas com “COVID-19” ou “SARS-CoV-2” no ambiente de trabalho. Os fatores de risco encontrados na literatura foram divididos em individuais e ocupacionais. Para esses autores, a relação positiva de idade e aumento do risco de mortalidade já está bem consolidada na literatura. Em uma série de estudos, indivíduos do sexo masculino apresentaram uma taxa de mortalidade superior ao sexo feminino. Essa diferença pode estar relacionada tanto a fatores genéticos quanto psicológicos e comportamentais. Devido a uma crença diferente sobre a severidade da COVID-19, por exemplo, homens tendem apresentar comportamentos mais

arriscados, baixas taxas de distanciamento social, uso de máscaras e busca de ajuda médica comparado às mulheres. Além desses fatores individuais, existem as comorbidades que estão diretamente associadas à saúde dos indivíduos.

No âmbito ocupacional, a natureza do trabalho tem sido reconhecida como um fator chave no contágio dos indivíduos. Diante da possibilidade de exposição ao vírus nos mais variados ambientes de trabalho. Os autores propõem uma classificação que vai de ambientes de trabalho com risco muito alto de exposição até alto, médio e baixo nível (LESO, FONTANA e IAVICOLI, 2021).

Zhang (2021) considera relevante o potencial de risco gerado pelo ambiente de trabalho, visando uma análise mais detalhada, o autor utilizou a base de dados (Occupational Information Network - O*NET) que contém informações que descrevem características das ocupações relacionadas ao risco de infecção por doenças contagiosas. Foram identificadas seis características com potencial de afetar o risco de contágio nos trabalhadores, são elas: contato com outros, local de trabalho apertado, duração semanal do trabalho, exposição a doenças e infecções, discussão face a face e proximidade física. Após uma análise estatística dessas variáveis, foi constatado que apenas o nível de exposição à doença e a proximidade física apresentaram correlação significativa com o contágio.

Marinaccio et al. (2020) também classificaram os setores de atividade em quatro níveis de risco: baixo, médio-baixo, médio-alto e alto. Essa classificação foi feita baseando-se na probabilidade de exposição, índice de proximidade e fator de agregação. Para os dois primeiros foi feita uma adaptação da metodologia utilizada pela O*Net americana. O último considera o nível de agregação social conectada com o trabalho. Esse tipo de caracterização do ambiente de trabalho, segundo os autores, é importante para a elaboração de uma estratégia de vigilância e vacinação mais efetiva.

De Negri et al. (2021) analisaram como fatores socioeconômicos afetam a probabilidade de morte de indivíduos que contraíram a COVID-19 no Estado do Rio de Janeiro. Para isso, os autores utilizaram um modelo de regressão logística com dados socioeconômicos do mercado de trabalho formal. A análise por ocupação e atividade econômica mostra que os trabalhadores de atividades essenciais que não podem trabalhar remotamente, tornam-se mais vulneráveis à COVID-19. Em especial, trabalhadores da saúde e da segurança pública apresentam risco maior de morte de 2,46 e 2,25 vezes, respectivamente. Além disso, indivíduos com escolaridade de nível superior apresentam um percentual de 44% menos chance de morrer em relação aos demais.

Conforme apresentado na literatura, a transmissão do vírus é um risco contínuo que pode ocorrer em qualquer local, não só limitado ao ambiente de trabalho. Sendo desafiador determinar quanto da transmissão do SARS-COV-2 ocorre nas firmas, e quanto está relacionado com as relações sociais, o ambiente doméstico ou exposição no transporte. O risco de contágio é uma complexa combinação entre fatores ambientais e humanos que estão associados com a probabilidade de infecção.

3. METODOLOGIA

3.1 Dados

O banco SUS_COVID disponibilizado pela Secretaria Estadual de Saúde de Pernambuco consiste em uma base de dados que contém informações sobre os casos de COVID-19 para o todo o estado de Pernambuco entre abril de 2020 e maio de 2021, além de informações sobre: resultado do teste, evolução do caso, sintomas, localização por CEP, sexo, idade, raça.

Os dados relativos a renda e mercado de trabalho foram obtidos no banco de dados da Relação Anual de Informações Sociais (RAIS) para o ano de 2019. A RAIS é um instrumento importante para a coleta de dados referentes ao mercado de trabalho formal no Brasil foi sendo instituída pelo Decreto nº 76.900, de 23/12/75. A periodicidade de coleta é anual e abrange todo território nacional. Algumas das variáveis disponíveis no banco são sexo, raça, idade, grau de instrução, mês de admissão, mês de desligamento, ocupação, remuneração média mensal e características do estabelecimento de trabalho.

A base de dados ESUS continha inicialmente 1.565.705 observações de indivíduos que realizaram o teste da COVID-19 para o estado de Pernambuco. Esses indivíduos foram georreferenciados no software *Qgis* a partir de suas coordenadas (latitude e longitude). As coordenadas foram obtidas a partir dos CEPs informados no momento do cadastro com a utilização de duas aplicações: Google Maps e Viacep. Desse modo, foi possível determinar em qual região da cidade do Recife estão localizados os indivíduos que realizaram os testes.

As duas bases (ESUS e RAIS) foram unidas utilizando-se o CPF dos indivíduos como identificador. A base de dados resultante passou a ter 252.397 observações, referentes aos indivíduos que realizaram os testes na cidade do Recife e são trabalhadores formais com registro na RAIS. Alguns indivíduos possuem observações duplicadas na base de dados, isso deve ao fato de terem realizado o teste mais de uma vez. Para estes, o resultado do teste da COVID-19 foi considerado positivo se pelo menos um resultado deu positivo e negativo caso contrário. Algumas variáveis estão presentes nos dois bancos como idade, raça e sexo. A escolha da variável para os modelos foi feita conforme aquela que estava melhor preenchida entre as duas bases de dados.

Da base de dados unificada foi extraída uma amostra que compreende apenas os trabalhadores do mercado formal da cidade do Recife que foram testados e que trabalham em firmas que possuem mais de 5 trabalhadores, correspondendo 29.943 observações. Foi necessário escolher um número mínimo de trabalhadores por firma, pois para firmas que possuem apenas um trabalhador não é possível calcular a proporção de infectados por firma.

A identificação da localidade dos trabalhadores foi uma parte importante do trabalho, pois os bairros foram usados como nível de agregação para a construção de algumas variáveis utilizadas nos modelos. Para identificar essas localidades foi utilizada a localização por CEP dos indivíduos no intuito gerar as coordenadas geográficas e o bairro ao qual estes pertencem.

Ambas bases de dados permitirão a realização de um estudo baseado em um conjunto de características socioeconômicas a nível de indivíduos. Até o momento, não são muitos os estudos sobre a COVID-19 que trazem esse nível de desagregação nos dados, o que permite uma contribuição importante para a literatura. Além disso, as observações da base de dados possuem uma dimensão espacial, que é de fundamental importância quando trabalhamos com contágio por vírus.

A descrição das variáveis encontra-se na tabela 2 da seção 4. A classificação dos setores de atividade foi feita com base no paper de De Negri et. al 2020. Os autores classificaram os distintos setores de atividade econômica entre setores essenciais e não essenciais com base no que foi definido pelas autoridades do estado do Rio de Janeiro.

3.3 Estratégia Empírica

Argumenta-se que determinadas características do ambiente de trabalho contribuem para a elevação da probabilidade de contágio entre os indivíduos de uma mesma firma. Conforme mencionado na literatura, os principais fatores que envolvem esse aumento de risco são: exposição geral ao público, proximidade física entre trabalhadores, possibilidade de conversão do trabalho presencial em remoto, dentre outros (MARINACCIO, 2020; ZHANG, 2021).

O modelo contém a dimensão espacial incorporada nas variáveis explicativas, pois a COVID-19 do indivíduo é explicada pelos contágios dos vizinhos de firma e de bairro. Isso implica que a variável de interesse é endógena, pois existe um processo de causalidade simultânea, onde os colegas impactam o indivíduo i e este, por sua vez, afeta

seus colegas de firma. Além disso, é sempre possível que outros fatores não observáveis possam estar afetando conjuntamente as variáveis X e Y. Isso reforça a importância de pensar em uma estratégia empírica que vise lidar com a endogeneidade do modelo.

A presença de endogeneidade na relação entre as variáveis ocasiona a estimação de parâmetros viesados para o modelo, caso não seja feita uma abordagem empírica adequada. Gibbons e Overman (2012) defendem que a identificação na econometria espacial é tão importante que deve ir para o centro do palco dos trabalhos aplicados.

Uma vez que a variável dependente é binária (valor = 1, caso o indivíduo tenha testado positivo e valor = 0, caso tenha testado negativo) é necessário utilizar um modelo de resposta binária para calcular o efeito marginal da variável de interesse. Devido a limitações no modelo de probabilidade linear para resposta binária é necessário pensar em modelos alternativos (GREENE, 2012). Em geral, modelos de índice de resposta binária são descritos da seguinte maneira:

$$P(COVID = 1 | \mathbf{x}) = F(\mathbf{x}\beta) \quad (1)$$

Onde $P(COVID = 1 | \mathbf{x})$ é a probabilidade de sucesso condicional ao vetor $1 \times K$ de variáveis explicativas \mathbf{X} , β é o vetor $K \times 1$ dos coeficientes correspondentes da regressão a ser estimada, e $F(*)$ (com $0 < F(z) < 1$ para todo $z \in \mathbb{R}$) denota uma função de distribuição acumulada (FDA), relacionando o índice $(\mathbf{x}\beta)$ à probabilidade.

Seguindo essa mesma lógica, a formulação matemática do modelo *probit* é dado por:

$$F(\mathbf{x}\beta) = \Phi(\mathbf{x}\beta) \equiv \int_{-\infty}^{\mathbf{x}\beta} \phi(z) dz, \quad (2)$$

Esse modelo é encontrado quando assumimos que u possui uma FDA Normal Padrão denotada por $\Phi(*)$, os parâmetros podem igualmente ser estimados pelo método da máxima verossimilhança.

O modelo empírico que será estimado está especificado abaixo, onde o coeficiente β_1 descreve o efeito da variável de interesse:

$$y_{ifk} = \beta_0 + \beta_1 P_COV_{ifk} + \beta_2 COV_BAIRRO_{ifk} + \sum_{n=2}^N \beta_n X_{nifk} + \varepsilon_{ifk} \quad (3)$$

onde:

y_{ifk} é uma variável binária igual a 1 se o indivíduo i que trabalhava na firma f no ano de 2019, e morava no bairro k , contraiu COVID-19 no ano de 2020; zero caso contrário.

$P_{COV_{ifk}}$ representa a proporção de casos positivos da COVID-19 dos colegas de trabalho do indivíduo i que mora no bairro k e trabalham na firma f .

$COV_{BAIRRO_{ifk}}$ refere-se a proporção de infectados no bairro k do indivíduo i que trabalhava na firma f .

X_{nifk} representa variáveis de controle como idade, sexo, raça, salário, escolaridade e setor de atividade.

ε_{ifk} corresponde ao erro aleatório do modelo

Para estimar o efeito causal da variável de interesse, exposta acima, a abordagem empírica adotada será a de variáveis instrumentais. Existem, basicamente, dois pressupostos que devem ser satisfeitos para que uma variável seja um bom instrumento em um modelo econométrico. A primeira é a exogeneidade, ou seja, a variável não deve estar correlacionada com o termo de erro. A segunda é a relevância, isto é, a variável deve estar correlacionada com a variável independente de interesse, X (GREENE, 2012).

Seguindo a mesma linha proposta por Kelejian e Robinson (2003), Reed (2015) e Feng e Humphreys (2016), que utilizam características dos vizinhos dos vizinhos como variável instrumental, o instrumento aqui utilizado refere-se à média de casos nos bairros dos colegas j da firma do indivíduo i , excluindo o bairro do próprio indivíduo i . O argumento para a exogeneidade deste instrumento é que a alta taxa de COVID nos bairros dos colegas j , afeta estes, mas não afeta o indivíduo i , pois o indivíduo i não habita estes bairros.

Primeiro estágio

$$P_{COV_{ifk}} = a_0 + a_1 W COV_{BAIRRO_{ifk}} + \sum_{n=3}^N \beta_n X_{nifk} + \varepsilon_{ifk} \quad (4)$$

Onde, W é a matriz quadrada $n \times n$ normalizada na linha, onde cada entrada relaciona os indivíduos à proporção dos vizinhos de bairro dos colegas de firma pegaram COVID-19, desde que não morem no mesmo bairro do indivíduo i . X_{nifk} são variáveis de controle como idade, sexo, raça, salário, escolaridade do indivíduo e setor de atividade da firma.

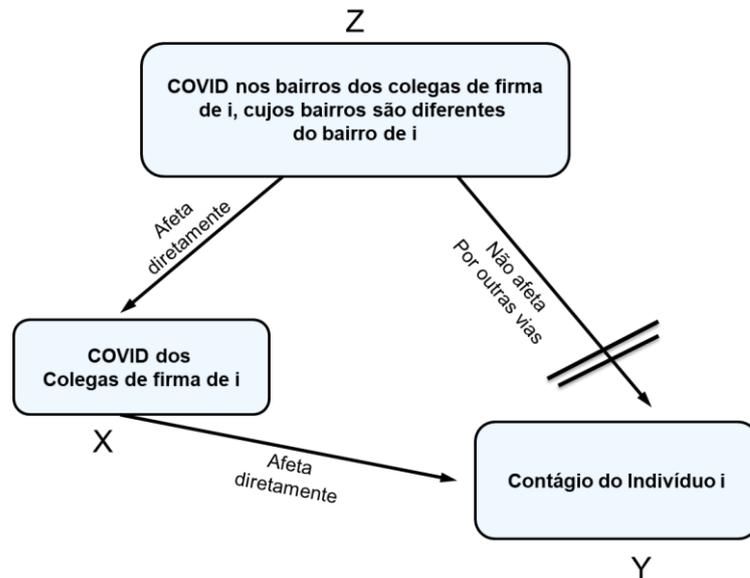
Segundo Estágio:

$$y_{ifk} = \beta_0 + \beta_1 P_COV_{ifk} + \beta_2 COV_BAIRRO_{ifk} + \sum_{n=3}^N \beta_n X_{nifk} + \varepsilon_{ifk} \quad (5)$$

Onde a variável P_COV_{ifk} representa a estimação no primeiro estágio da proporção de COVID-19 nas firmas. As demais variáveis são as mesmas variáveis de controle elencadas anteriormente.

O diagrama abaixo ajuda a esclarecer a relação entre as principais variáveis do modelo:

Figura 1 – Diagrama ilustrativo do instrumento utilizado



Fonte: Elaboração Própria.

Como ilustrado na figura acima, o colega j de mesmo bairro do indivíduo i não é considerado para o instrumento, pois a COVID no bairro desse colega j o afeta, mas também afeta i, o que invalidaria o instrumento.

Assim, conseguimos uma média das proporções da COVID-19 nos bairros dos colegas, excluindo o bairro do próprio indivíduo i.

4. RESULTADOS

4.1 Análise descritiva

Na tabela a seguir temos um resumo descritivo das variáveis utilizadas.

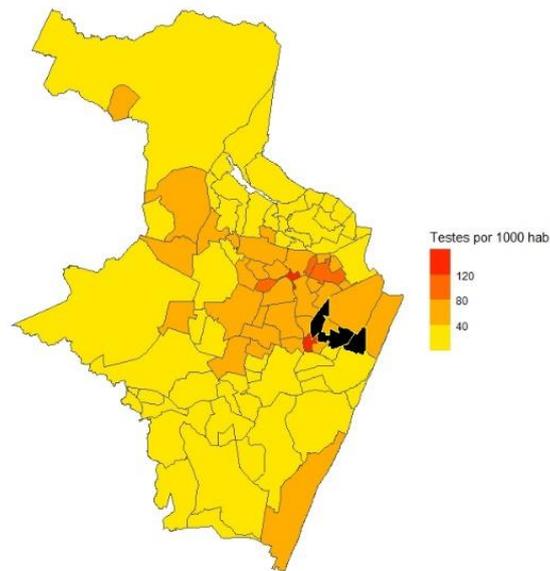
Tabela 2 – Resumo descritivo das variáveis consideradas no modelo

Variável	Média	Desvio Padrão	Min	Max
Covid	0,334	0,472	0	1
P_Cov	3,516	3,337	0	100
Cov_bairro	29,297	2,988	9,09	61,54
W_Cov_bairro	29,329	0,905	20,04	43,80
Sexo	0,440	0,496	0	1
Idade	46,891	12,664	15	79
Log da renda	4916,154	6654,660	0	99660,28
Pretos e pardos	0,466	0,499	0	1
Deficiência	0,017	0,128	0	1
Fundamental	0,057	0,233	0	1
Médio	0,432	0,495	0	1
Superior	0,464	0,499	0	1

A tabela acima resume as variáveis que serão utilizadas no modelo. A variável sexo é uma *dummy* = 1 para o sexo masculino e 0 para feminino. Pretos e pardos é uma variável *dummy* = 1 para pessoas pretas ou pardas e 0, caso contrário. A escolaridade é representada pelas variáveis *dummies* fundamental, médio e superior que corresponde aos indivíduos que concluíram estas etapas de ensino. Log da renda é o logaritmo da renda mensal dos indivíduos conforme a RAIS. Deficiência é uma variável *dummy* indicando se o trabalhador é portador de deficiência.

O mapa a seguir mostra a distribuição da COVID-19 nos bairros de Recife, considerando o número de infectados sobre o número de testados.

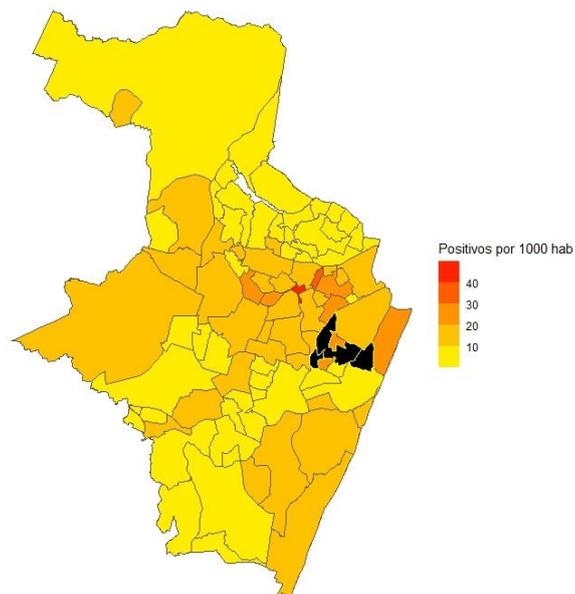
Figura 2 – Mapa número de testes realizados por bairros de Recife por 1000 habitantes.



Obs: Os bairros do centro de Recife estão em preto, pois apesar de serem pouco habitados são muito frequentados isso torna a taxa de COVID-19 altíssima (*outliers*) distorcendo os valores para os demais.

Analisando a figura acima, percebemos que a maior parte dos bairros do Recife apresentam até 120 testes por 1000 habitantes, apenas 3 deles se destacam do restante e estão preenchidos com a cor preta no mapa, são eles: Santo Antônio, Boa Vista e Paissandu onde foram realizados 452,83; 362,23; e 130,18 testes por 1000 habitantes, respectivamente.

Figura 3 – Mapa da proporção de infectados sobre a população de cada bairro em Recife



A maior parte dos bairros registraram até 40 casos positivos por 1000 habitantes. Os bairros que apresentaram a maior ocorrência estão preenchidos na cor preta, foram eles: Santo Antônio, Boa Vista e Paissandu com 116,98; 93,65 e 78,89 respectivamente.

Na tabela abaixo, podemos observar a distribuição da população e dos casos da COVID-19 por sexo. O conjunto dos trabalhadores do sexo feminino corresponde a 58,43% do total e 32,13% apresentaram teste positivo para COVID-19, enquanto na população masculina esse número é levemente superior, 34,68%.

Tabela 3 – Casos de COVID-19 para os trabalhadores do Recife por sexo

Sexo	Casos Confirmados	Trabalhadores Testados	(%)
Feminino	5374	16724	32.13
Masculino	4584	13219	34.68
Total	9958	29943	33.26

Fonte: RAIS e SES/PE

Na variável raça, a maior população dentre os grupos é a de pardos, 12268 indivíduos (52,38%), seguida dos brancos 8179 (36,27%), a menor população da amostra considerada é de indígenas, composta de 16 indivíduos. O percentual de contágio fica em torno de 33% para a maioria das raças com exceção dos indígenas, que apresentam um percentual menor de infecção, 6,25%.

Tabela 4 – Casos de COVID-19 para os trabalhadores do Recife por raça

Raça	Casos Confirmados	Trabalhadores Testados	(%)
Amarela	418	1208	34.60
Branca	2735	8179	33.44
Ignorado	2178	6772	32.16
Índigena	1	16	6.25
Parda	4132	12268	33.68
Preta	494	1500	32.93
Total	9958	29943	33.26

Fonte: RAIS e SES/PE

Analisando os casos por faixa etária, percebemos que os distintos grupos sofrem contágio em proporções semelhantes. O grupo mais afetado é de idosos entre 71 e 100 anos com 38,97% de casos, enquanto o menos afetado é o grupo de trabalhadores que possuem até 20 anos com 28,82%.

Tabela 5 – Casos de COVID-19 para os trabalhadores do Recife por faixa etária

Faixa Etária	Casos Confirmados	Trabalhadores Testados	(%)
0-20	140	485	28.82
21-30	1447	4853	29.82
31-40	186	518	35.91
41-50	3280	9205	35.63
51-60	3658	11089	32.99
61-70	1169	3593	32.54
71-100	78	200	38.97
Total	9958	29943	33.26

Fonte: RAIS e SES/PE

Também é importante analisar nos dados como está a taxa de contágio entre os setores de atividade econômica, pois cada setor possui características específicas que o tornam mais ou menos propenso à transmissão da COVID-19.

Na tabela 6, podemos observar que o setor que possui menor percentual de casos é o de Imprensa e Informação, 29,26%. Isso é algo esperado para o setor, pois os trabalhadores, em geral, guardam certo distanciamento entre si, o que dificulta a transmissão do vírus que ocorre principalmente pelo contato físico. O setor de petróleo gás e biocombustíveis apresenta o maior percentual de casos para esta amostra de trabalhadores com 53,70%. Outros setores que apresentam uma taxa de contágio acima da média (29,29%) são Serviço de Saúde e Serviço de Utilidade Pública com 40,70% e 40,64%, respectivamente.

Tabela 6 – Casos de COVID-19 para os trabalhadores do Recife por setor de atividade

Setor	Casos Confirmados	Trabalhadores Testados	(%)
Agropecuária e Pesca	17	53	32.08
Alimentação Banco Jurídica etc	602	1703	35.35
Comercio Essencial	403	1261	31.96
Comercio NE *	219	667	32.83
Construção	186	633	29.38
Fabricação produtos essenciais	135	368	36.68
Imprensa e Informação etc	328	1121	29.26
Petro Gás e Biocombustíveis	29	54	53.70
Prod Florestal e Indústria NE*	137	382	35.86
Seg e Def Civil	271	733	36.97
Serv Publico	3892	12063	32.26
Serv Saúde	1285	3157	40.70
Serv Utilidade Pública	139	342	40.64
Serviços NE	1797	5872	30.60
Transporte de cargas	409	1091	37.49
Total	9958	29943	33.26

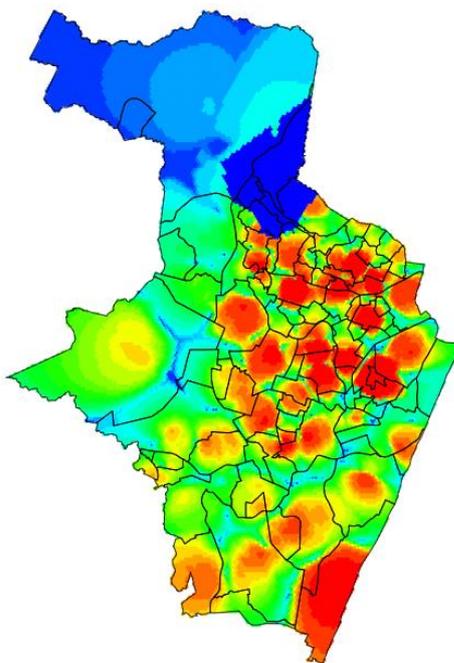
Fonte: RAIS e SES/PE. Nota: * NE são atividades não essenciais

4.2 Resultados dos modelos e Discussão

A análise que está sendo conduzida neste trabalho refere-se a eventos pontuais (ocorrências da COVID-19) que estão distribuídos espacialmente. Neste sentido, é importante verificar se a distribuição espacial se dá de forma aleatória ou se há algum padrão de correlação espacial entre os pontos. Carvalho e Câmara (2004) consideram que para este tipo de análise pode ser utilizada a estimação da função *Kernel*, que é descrita como uma alternativa para analisar o comportamento de padrões de pontos em uma região de estudo. Isso é feito por meio de uma função bi-dimensional que realiza a contagem de todos os pontos de uma região e pondera pela distância de cada um à localização de interesse.

O mapa abaixo descreve a estimação da função *kernel* para a cidade do Recife.

Figura 4 - Mapa Kernel da incidência dos casos de COVID-19 nos bairros de Recife



Analisando o mapa podemos perceber que há uma incidência maior dos casos da doença nos bairros da região central da cidade com uma diminuição da intensidade conforme se avança para as extremidades. Os bairros da região Norte e Oeste são os menos afetados, enquanto os bairros da região centro-leste e Sul do município têm uma maior incidência.

A seguir serão apresentados os resultados obtidos pelos modelos de probabilidade de contágio da COVID-19 para os indivíduos da cidade do Recife. Faz-se necessário

ressaltar que os resultados apresentados serão subdivididos em estratos por quantidade de trabalhadores em uma mesma firma, pois é razoável pensar que em firmas com maior número de funcionários o “efeito aglomeração” potencializa o contágio.

Na tabela abaixo são apresentados os modelos *Probit*, o primeiro e segundo estágio do *IV Probit* para firmas que possuem pelo menos 5 trabalhadores.

Tabela 7 - Regressões para firmas com 5 ou mais trabalhadores. (Efeitos Marginais)

	Probit		1º Estagio		IV Probit	
	Coef.	p-valor	Coef.	p-valor	Coef.	p-valor
P_Cov	0,0050	0,000	-	-	0,0220	0,003
W_Cov_bairro	-	-	0,4182	0,000	-	-
Cov_bairro	0,0061	0,000	0,0326	0,000	0,0057	0,000
Sexo	0,0271	0,000	-0,4401	0,000	0,0343	0,000
Pretos e pardos	0,0088	0,117	0,0338	0,371	0,0085	0,131
Idade	0,0069	0,000	0,0100	0,284	0,0067	0,000
Idade2*	-0,0001	0,000	0,0002	0,119	-0,0001	0,000
Log da renda	0,0058	0,000	0,0626	0,000	0,0044	0,010
Deficiência	0,0249	0,242	-0,6664	0,000	0,0344	0,111
Fundamental	-0,0261	0,131	-0,1055	0,362	-0,0238	0,169
Médio	-0,0070	0,611	0,1740	0,059	-0,0103	0,453
Superior	-0,0152	0,273	0,2488	0,008	-0,0207	0,139
Obs	29481		29481		29481	
Wald Test Exogeneity	-		-		p = 0,0210	
Teste F	-		129,57		-	

Nota: Os valores apresentados são os efeitos marginais das regressões. * Idade2 é a idade ao quadrado.

Primeiramente, a seguir, vamos descrever as variáveis utilizadas nos modelos:

No resultado apresentado podemos observar, como já esperado, que a proporção de COVID-19 na firma (**P_Cov**) é estatisticamente significativa e positivamente correlacionada com a variável dependente (contágio da COVID-19 pelo indivíduo). O efeito marginal para o modelo *Probit* indica que um aumento em um ponto percentual dos infectados nas firmas com 5 ou mais trabalhadores corresponde a um incremento de cerca 0,5% na probabilidade de contágio do indivíduo. No modelo *IV Probit*, temos que o aumento de 1 ponto percentual no número de infectados das firmas, implicará em um incremento de cerca 2,2% na probabilidade de contágio do indivíduo.

É pré-requisito para o uso de variáveis instrumentais que haja correlação positiva entre o instrumento e a variável de interesse. O resultado do primeiro estágio, na tabela acima, mostra que há correlação positiva (coeficiente 0,418) e estatisticamente

significante entre o instrumento (**W_Cov_bairro**) e a variável de interesse (**P_Cov**). O resultado é esperado, pois quanto mais contágio no bairro de um indivíduo maior a chance de infecção. Testes adicionais podem ser empregados para testar a validade de instrumento.

Na parte inferior da tabela 2, temos o teste de Wald e o teste F, ambos podem ser utilizados para testar a validade do instrumento. O teste de Wald possui como hipótese nula a não endogeneidade, o resultado do teste para o modelo acima ($p = 0,0210$) rejeita a hipótese nula, portanto o modelo é endógeno, o que justifica o uso de instrumento. O teste F com valor inferior a 10 indica instrumentos fracos segundo Greene (2012), o resultado desse teste para o modelo apresentado (129,57) é largamente superior como pode ser visto na tabela acima.

Para verificar se os resultados continuam válidos para firmas de porte maior, será apresentado na tabela seguinte as regressões para firmas com mais de 50 trabalhadores:

Tabela 8 - Regressões para firmas com 50 ou mais trabalhadores

	Probit		1º Estagio		IV Probit	
	Coef.	p-valor	Coef.	p-valor	Coef.	p-valor
P_Cov	0,0043	0,000	-	-	0,0254	0,004
W_Cov_bairro	-	-	0,3908	0,000	-	-
Cov_bairro	0,0063	0,000	0,0239	0,000	0,0060	0,000
sexo	0,0261	0,000	-0,3990	0,000	0,0343	0,000
Pretos e pardos	0,0113	0,051	0,0402	0,145	0,0107	0,064
Idade	0,0069	0,000	0,0086	0,208	0,0067	0,000
Idade2*	-0,0001	0,000	0,0002	0,005	-0,0001	0,000
Log da renda	0,0062	0,000	0,0362	0,000	0,0052	0,003
Deficiência	0,0190	0,381	-0,4993	0,000	0,0277	0,206
Fundamental	-0,0223	0,206	-0,0797	0,343	-0,0204	0,248
Médio	-0,0029	0,832	0,0576	0,387	-0,0048	0,733
Superior	-0,0129	0,360	0,1036	0,124	-0,0167	0,237
Obs	27922		27922		27922	
Wald Test Exogeneity	-		-		p = 0.0164	
Teste F	-		312,65		-	

Neste resultado também temos que (**P_Cov**) é estatisticamente significante e positivamente correlacionado com a variável dependente. O efeito marginal para o modelo *Probit* indica que um aumento em um ponto percentual dos infectados nas firmas com 50 ou mais trabalhadores corresponde a um incremento de cerca 0,43% na probabilidade de contágio do indivíduo. No modelo *IV Probit*, temos que o aumento de 1

ponto percentual no número de infectados das firmas, implicará em um incremento de cerca de 2,5% na probabilidade de contágio do indivíduo.

No primeiro estágio também temos uma relação positiva (0,391) entre o instrumento e a variável de interesse que é estatisticamente significativa a 1%. O modelo também é endógeno conforme o teste de exogeneidade de *Wald* ($p = 0,0210$) e o teste F ($312,65 > 10$) rejeita a hipótese nula de que o instrumento empregado seja fraco.

Na tabela 4, os coeficientes da variável de interesse mantêm os mesmos sinais das regressões anteriores para os três modelos e permanecem estatisticamente significantes.

Tabela 9 - Regressões para firmas com 200 ou mais trabalhadores

	Probit		1° Estagio		IV Probit	
	Coef.	p-valor	Coef.	p-valor	Coef.	p-valor
P_Cov	0,0041	0,003	-	-	0,0236	0,004
W_Cov_bairro	-	-	0,5257	0,000	-	-
Cov_bairro	0,0062	0,000	0,0288	0,000	0,0059	0,000
sexo	0,0276	0,000	-0,4609	0,000	0,0364	0,000
Pretos e pardos	0,0095	0,129	0,0539	0,065	0,0088	0,160
Idade	0,0071	0,000	0,0238	0,001	0,0066	0,000
Idade2*	-0,0001	0,000	0,0001	0,235	-0,0001	0,000
Log da renda	0,0068	0,000	0,0141	0,099	0,0061	0,001
Deficiência	0,0130	0,580	-0,6134	0,000	0,0224	0,346
Fundamental	-0,0239	0,204	-0,1738	0,048	-0,0206	0,274
Médio	-0,0070	0,631	-0,0106	0,878	-0,0075	0,609
Superior	-0,0233	0,117	-0,0192	0,784	-0,0247	0,095
Obs	23973		23973		23793	
Wald Test Exogeneity	-		-		p = 0,0179	
Teste F	-		325,34		-	

Neste caso, o aumento de 1 ponto percentual no modelo sem instrumento corresponde a um incremento de 0,41% no risco de contágio pelo indivíduo. No modelo com instrumento, esse valor é cerca de 5 vezes superior, levando a um incremento de 2,36% no risco de contágio.

Em relação as variáveis individuais e socioeconômicas: a *dummy* de sexo, a idade e o salário também apresentam significância estatística a pelo 5% nos modelos. Por outro lado, ambos os modelos não apresentam significância estatística para a raça de pretos e pardos, deficientes e para os níveis de escolaridade.

No que se refere a idade e sexo, os resultados estão compatíveis com os encontrados em Carlsten et. al (2021), os autores apontam que pessoas mais velhas e do sexo masculino encontram-se mais vulneráveis tanto ao contágio como morte pelo vírus.

A despeito de não ter sido encontrado um resultado estatisticamente significativo para raça, na literatura os trabalhos de Quan et al. (2021), Mutambudzi et al. (2021), De Negri et al. (2021) dentre outros apontam para um maior nível de contágio e/ou mortes para indivíduos pretos ou pardos.

A partir dos resultados encontrados podemos concluir que o ambiente de trabalho contribui significativamente para a elevação no risco de contágio entre os trabalhadores. Esses resultados devem reforçar a importância das autoridades na elaboração de políticas que possam garantir um ambiente mais seguro para os trabalhadores. Devido a relevância do ambiente de trabalho para o contágio dos indivíduos, autores como Carlsten et al. (2021) sugerem o tratamento da COVID-19 como doença ocupacional.

Por esse motivo, algumas medidas preventivas já propostas na literatura podem ser tomadas, tais como diretivas de trabalho remoto, acesso ao local de trabalho restrito aos funcionários essenciais, distanciamento físico, triagem de rotina, isolamento de pessoas infectadas, rastreamento e quarentena de contatos, higiene das mãos, monitoramento ambiental e uso adequado de equipamentos de proteção individual e vacinas dentre outras medidas baseadas em evidências (CARLSTEN, 2021; LESO FONTANA e IAVICOLI, 2021; OMS e OIT, 2021).

Para a implementação dessas políticas, é importante que o empregador entenda o nível de risco ao qual estão submetidos seus funcionários. No mesmo relatório a OMS sugere a classificação dos ambientes de trabalho em três níveis: baixo risco de exposição, médio risco de exposição e alto risco de exposição. No primeiro, as tarefas não teriam um contato com frequente com o público geral e internamente entre os próprios colegas de trabalho. No segundo nível de risco, estariam os estabelecimentos que possuem um contato frequente tanto do público interno quanto externo. No terceiro nível, há um contato próximo entre pessoas com a Covid-19 ou suspeita, gerando um altíssimo risco de transmissão, o exemplo mais comum para esse caso é o setor de saúde que trabalha diretamente com pessoas infectadas (OMS e OIT, 2021).

4.3 Testes de Robustez

Nesta seção serão realizados testes de robustez com a finalidade de verificar se os resultados apresentados anteriormente são consistentes. O primeiro teste consiste na aleatorização da variável de interesse do modelo *Probit*, proporção de COVID-19 na firma. Esse procedimento atribui para as observações valores aleatórios para a variável de interesse, permitindo testar se a variável apresenta um efeito significativo de fato ou apenas uma correlação espúria e acidental. Em seguida, será feita uma aleatorização na variável instrumental do modelo IV Probit.

A tabela a seguir apresenta o teste para os modelos com e sem variável instrumental em firmas com 5 ou mais trabalhadores.

Tabela 10 – Teste de Robustez para modelo com 5 ou mais trabalhadores

	Probit		1º Estagio		IV Probit	
	Coef.	p-valor	Coef.	p-valor	Coef.	p-valor
P_Cov	0.0006	0.531	-	-	-0.0145	0.691
W_Cov_bairro	-	-	0.1291	0.000	-	-
Cov_bairro	0.6318	0.000	2.3964	0.000	0.6600	0.000
sexo	0.0252	0.000	-0.7351	0.000	0.0143	0.616
Pretos e pardos	0.0088	0.115	0.0556	0.152	0.0096	0.100
Idade	0.0074	0.000	0.0157	0.098	0.0075	0.000
Idade2*	-0.0001	0.000	0.0002	0.071	-0.0001	0.000
Log da renda	0.0055	0.000	0.1045	0.000	0.0070	0.075
Deficiência	0.0459	0.029	-0.4894	0.001	0.0385	0.179
Fundamental	-0.0227	0.188	-0.0605	0.611	-0.0231	0.182
Médio	0.0031	0.822	0.3234	0.001	0.0078	0.661
Superior	-0.0137	0.318	0.5383	0.000	-0.0057	0.813
Obs	29481		29481		29481	
Wald Test Exogeneity	-		-		p = 0,5950	
Teste F	-		55,43		-	

No primeiro modelo, com aleatorização da variável de interesse, não observamos uma relação estatisticamente significativa da variável **P_Cov** com a variável dependente do modelo. O mesmo resultado ocorre no modelo IV probit quando a variável aleatorizada é o instrumento, nesse caso, apesar de o instrumento ser estatisticamente significativo no

primeiro estágio, não há uma relação significativa entre a variável de interesse e a variável dependente do modelo.

Tabela 11 – Robustez para o modelo de 200 ou mais trabalhadores

	Probit		1º Estagio		IV Probit	
	Coef.	p-valor	Coef.	p-valor	Coef.	p-valor
P_Cov	0.0006	0.531	-	-	-0.0852	0.177
W_Cov_bairro	-	-	0.0877	0.030	-	-
Cov_bairro	0.6381	0.000	0,0153	0.004	0.6717	0.000
sexo	0.0245	0.000	-0.8542	0.000	-0.0521	0.386
Pretos e pardos	0.0091	0.144	0.0792	0.013	0.0144	0.029
Idade	0.0075	0.000	0.0465	0.000	0.0102	0.000
Idade2*	-0.0001	0.000	0.0000	0.881	-0.0001	0.013
Log da renda	0.0065	0.000	0.1154	0.000	0.0154	0.011
Deficiência	0.0379	0.102	-0.3802	0.002	0.0001	0.997
Fundamental	-0.0210	0.265	-0.1096	0.252	-0.0268	0.139
Médio	0.0031	0.829	0.1906	0.011	0.0188	0.292
Superior	-0.0216	0.142	0.4421	0.000	0.0190	0.593
Obs	23973		23973		23973	
Wald Test Exogeneity	-		-		p= 0.2438	
Teste F	-		123,12			

O mesmo teste foi feito para o modelo com 200 ou mais tabalhadores por firma. Nos modelo Probit e IV Probit o resultado permanece semelhante ao encontrado anteriormente. Interessante notar também que no teste de Wald nas duas tabelas, a hipótese nula de não endogeneidade não é rejeitada. Os testes realizados permitem afastar suspeitas de que os resultados encontrados são apenas fruto de uma correlação espúria entre as variáveis ou do acaso.

5. CONSIDERAÇÕES FINAIS

Centenas de milhões de pessoas foram infectadas, e outras tantas morreram, pelo vírus da COVID-19 em todo mundo. Em razão da relevância e urgência deste assunto muitos na academia, governos e muitas outras organizações dedicaram esforços para entender como o vírus se propaga e afeta os infectados. Nesse contexto, este trabalho buscou compreender o impacto que o ambiente de trabalho exerce no contágio analisando como a proporção de indivíduos contaminados numa firma e as características socioeconômicas destes afetam a probabilidade de contágio pelo vírus da COVID-19.

Os resultados aqui apresentados apontam que firmas que possuem mais trabalhadores contaminados, tornam o indivíduo mais propenso a também ser contaminado. Porém, o mais interessante é que o impacto que as firmas exercem na probabilidade de contágio é muito alto, sendo até superior ao contágio gerado pelo bairro em que os indivíduos residem.

Diante desse resultado é importante que os órgãos competentes formulem estratégias que contemplem o ambiente de trabalho dos indivíduos. Atualmente, existem cartilhas, relatório e estudos conduzidos por organizações como OIT e OMS, que oferecem algumas medidas que podem ajudar a mitigar o risco de doenças contagiosas.

Duas ressalvas importantes merecem ser pontuadas a respeito dos dados sobre os quais se baseiam os resultados. Primeiro, a amostra da RAIS contempla apenas trabalhadores do setor formal, isso implica que os indivíduos analisados estão em melhores condições socioeconômicas do que a média dos trabalhadores. Segundo, as informações de contágio estão sujeitas ao viés amostral daqueles que foram realizar os testes de detecção do vírus. Desse modo, sabemos que pessoas mais escolarizadas, de mais alta renda e trabalhadores da saúde, por exemplo, tendem a testar-se mais do que a população geral, gerando assim um viés de mais casos para esses grupos.

Portanto, esse estudo contribui com a literatura da transmissão da COVID-19 ao mostrar a importância que a firma exerce no contágio dos trabalhadores através de uma análise baseada em microdados, que têm sido pouco utilizados em estudos semelhantes na maioria dos casos por indisponibilidade de dados.

REFERÊNCIAS

AGUIAR, Bianca Fontana; SARQUIS, Leila Maria Mansano; MIRANDA, Fernanda Moura D.'Almeida. Sequelas da Covid-19: uma reflexão sobre os impactos na saúde do trabalhador. **Research, Society and Development**, São Paulo, v. 10, n. 14, p. e40101421886, out. 2021.

BRASIL. Ministério do Trabalho e Emprego. **O que é RAIS**. Brasília. Disponível em: <<http://pdet.mte.gov.br/o-que-e-rais>>. Acesso em: 17 de jun. de 2018.

BERKMAN, Lisa F.; KAWACHI, Ichirō; GLYMOUR, M. Maria (Ed.). **Social epidemiology**. Second Edition. Oxford: Oxford University Press, 2014. 614 p.

CARLSTEN, Christopher et al. COVID-19 as an occupational disease. **American Journal of Industrial Medicine**, New York, v. 64, n. 4, p. 227-237, Abr. 2021.

Carvalho, S. S. de; Nogueira, M. O. O trabalho precário e a pandemia: os grupos de risco na economia do trabalho. **Mercado de Trabalho: conjuntura e análise**, Rio de Janeiro. ano 26, n. 70, p. 49-68, set. 2020. Disponível em: <<https://bit.ly/3HgQZNp>>

CONSOLAZIO, David et al. Assessing the Impact of Individual Characteristics and Neighborhood Socioeconomic Status During the COVID-19 Pandemic in the Provinces of Milan and Lodi. **International Journal of Health Services, Cidade**, vol. 51, n. 3, p. 311-324, Jul. 2021.

Corseuil, C. H.; Ramos, L.; Russo, F. A evolução do emprego setorial em 2020: quão heterogêneo foi o tombo entre os setores? **Carta de Conjuntura**, Rio de Janeiro, n. 48, p. 1-16. 6 jul. 2020. Disponível em: <<https://bit.ly/3xtgszA>>.

Costa, J. S.; Barbosa, A. L. N. de H.; Hecksher, M. Desigualdades no mercado de trabalho e pandemia da covid-19. Rio de Janeiro: Ipea, ago. 2021. (Texto para Discussão, n. 2684). Disponível em: <<https://bit.ly/3MKYXzl>>

DE NEGRI, F. et al. Socioeconomic factors and the probability of death by Covid-19 in Brazil. **Journal of Public Health**, Oxford, vol. 43, n. 3, p. 493-498, Set. 2021.

DRUCK, S.; Carvalho, M.S.; Câmara, G.; Monteiro, A.V.M. "Análise Espacial de Dados Geográficos". In: Carvalho, M.S.; Câmara, G (Org). **Análise Espacial De Eventos**. Brasília: EMBRAPA, 2004. Cap. 2.

Feng, X., & Humphreys, B. Assessing the Economic Impact of Sports Facilities on Residential Property Values. **Journal of Sports Economics**, Morgantown, vol. 19, n. 2, p. 188-210, 2016.

FINLAY, Keith; MAGNUSSON, L. M.; SCHAFFER, M. E. weakiv: Weak-instrument-robust tests and confidence intervals for instrumental-variable (IV) estimation of linear, probit and tobit models. **Statistical Software Components**. Boston College Department of Economics, Out 2013.

GIBBONS, Stephen; OVERMAN, Henry G. Mostly pointless spatial econometrics?. **Journal of regional Science**, Hoboken, v. 52, n. 2, p. 172-191, 2012.

GREENE, W. H. **Econometric analysis**. 7th (International) ed. New York: Pearson, 2012. 978 p.

HECKSHER, M.; CORSEUIL, C. H. Mais empregos com menos aglomeração: proposta de desoneração da folha com ênfase nas menores jornadas. **Mercado de Trabalho: conjuntura e análise**, Rio de Janeiro. ano 26, n. 69, p. 1-11, Jul. 2020. Disponível em: <<https://bit.ly/34pe9Q7>>.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Pesquisa nacional de saúde 2019. Informações sobre domicílios, acesso e utilização dos serviços de saúde: Brasil, grandes regiões e unidades da federação. Rio de Janeiro: IBGE; 2020.

Kelejian, H. H., & Robinson, D. P. (1993). A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model. **Papers in Regional Science**, v. 72, n. 3, p. 297-312, 1993.

KING, Gary; ZENG, Langche. Logistic regression in rare events data. **Political analysis**, v. 9, n. 2, p. 137-163, 2001.

LESO, Veruscka; FONTANA, Luca; IAVICOLI, Ivo. Susceptibility to coronavirus (COVID-19) in occupational settings: The complex interplay between individual and workplace factors. **International Journal of Environmental Research and Public Health**, v. 18, n. 3, p. 1030, 2021.

Link BG, Phelan J. Social conditions as fundamental causes of disease. **Journal of health and social behavior**, p. 80-94, 1995.

MARMOT, Michael. Social determinants of health inequalities. **The lancet**, v. 365, n. 9464, p. 1099-1104, 2005.

MENA, Gonzalo E. et al. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. **Science**, v. 372, n. 6545, p. 1-8, 2021.

MUTAMBUDZI, Miriam et al. Occupation and risk of severe COVID-19: prospective cohort study of 120 075 UK Biobank participants. **Occupational and Environmental Medicine**, v. 78, n. 5, p. 307-314, 2021.

PIRES, Luiza Nassif; CARVALHO, Laura; XAVIER, Laura de Lima. COVID-19 e desigualdade: a distribuição dos fatores de risco no Brasil. **Experiment Findings**, v. 21, 2020.

Procópio, D. P. (2020). Fatores socioeconômicos associados à expansão de casos da COVID-19 no Rio Grande do Sul. **VITTALLE - Revista De Ciências Da Saúde**, v. 32, n. 3, p. 19-29, 2020.

QUAN, Daniel et al. Impact of race and socioeconomic status on outcomes in patients hospitalized with COVID-19. **Journal of general internal medicine**, v. 36, p. 1302-1309, Jan. 2021.

RUTTER, Paul D. et al. Socio-economic disparities in mortality due to pandemic influenza in England. **International journal of public health**, v. 57, n. 4, p. 745-750, 2012.

REED, William Robert. On the practice of lagging variables to avoid simultaneity. **Oxford Bulletin of Economics and Statistics**, v. 77, n. 6, p. 897-905, 2015.

ORGANIZAÇÃO MUNDIAL DA SAÚDE & ORGANIZAÇÃO INTERNACIONAL DO TRABALHO. Prevenção e mitigação da transmissão da COVID-19 no trabalho. Organização Mundial da Saúde; 2021.

ZHANG, Michael. Estimation of differential occupational risk of COVID-19 by comparing risk factors with case data by occupational group. American journal of industrial medicine, v. 64, n. 1, p. 39-47, 2021.