

Relatório Provisório da Pesquisa: Impactos da Pandemia na Ocupação dos Jovens no Mercado de Trabalho brasileiro

1. Introdução

Este relatório documenta, passo a passo, o desenvolvimento da presente pesquisa, que objetiva, em linhas gerais, mensurar o impacto do choque econômico no contexto da pandemia de COVID-19 sobre a ocupação juvenil no mercado de trabalho. Mais propriamente, dentro do referencial teórico do Modelo de Search & Matching de Diamond-Mortensen-Pissarides (DMP), estimar as probabilidades condicionadas (Cadeias de Markov) de transição entre emprego e desemprego, com foco na integração dos jovens no setor laboral, e complementação com a mensuração de heterogeneidades no impacto dado diferentes características individuais (como sexo, raça, escolaridade, localidade, experiência), mas sobretudo, aplicação do referido modelo para regressão da eficiência de matching para o público-alvo.

Para este fim, utilizou-se microdados da PNAD Contínua Trimestral (2019-2022) e dados de admissões formais disponibilizados pelo CAGED como proxy para o estoque de vagas.

2. Construção do Painel Longitudinal da PNADC

2.1. Obtenção dos Dados

Foram utilizados microdados da PNAD Contínua Trimestral, para os anos de 2019 a 2022, coletados a partir do módulo *datazoom_pnadcontinua* do STATA, desenvolvido pelo projeto Data Zoom da PUC Rio, que permite o emparelhamento dos indivíduos consultados pelo IBGE a partir da técnica proposta por Ribas e Soares (2008). Assim, às variáveis padrão são adicionados novos elementos de identificação *hous_id* e *ind_id* para identificação do domicílio e do indivíduo, respectivamente, caso sejam sorteados novamente na nova amostra. Estas informações serão úteis para a observação do mesmo indivíduo em dois ou mais trimestres consecutivos e registro de sua evolução no mercado de trabalho, para posterior geração das matrizes de transição e regressões contra as suas características e os períodos de análise, especialmente o que antecede e o em que se vive a pandemia.

O módulo aplicado às planilhas trimestrais da PNADC, baixadas diretamente do portal do IBGE (<https://www.ibge.gov.br/estatisticas/sociais/trabalho/9171-pesquisa-nacional-por-amostra-de-domicilios-continua-mensal.html?=&t=microdados>), retornou quatro arquivos de

dados .dta para cada ano de análise. Estes foram, em seguida, importados para o R, onde foram unidos em um mesmo banco de dados, via empilhamento, e feito o filtro das variáveis relevantes, conforme o script *importação_filtro.R*, tratado a seguir.

2.2. Filtragem dos Dados

No processo de filtro dos dados da PNADC já limpos e emparelhados pelo módulo anterior, prosseguimos para a seleção apenas das variáveis relevantes para o método desta pesquisa. Seguindo o convencionado pelo IBGE e estudos da OIT sobre mercado de trabalho, consideramos como Jovens apenas indivíduos na faixa etária de 14 a 29 anos de idade, e filtramos, restringindo às observações cuja variável de código *V2009* (que indica a idade relatada na semana de referência) se situassem neste intervalo de idades.

Para identificação do indivíduo, além das variáveis de emparelhamento geradas pela ferramenta do Data Zoom (*ind_id* foi renomeada para *id_pessoa*), e da idade, selecionamos variáveis idiossincráticas que podem ser úteis para análise dos impactos heterogêneos entre os mesmos jovens, como:

- *V2007*: Sexo (se homem, 1; se mulher, 2)
- *V1022*: Situação do domicílio (se urbana, 1; se rural, 2)
- *V2010*: Cor ou raça (branca, 1; preta, 2; parda, 4; e outras)
- *V3009A*: Escolaridade (fundamental ou menos, 1-3; médio, 4-5; superior, 6)

Seguindo a literatura que compõem o referencial teórico da pesquisa, consideramos a questão da acumulação de experiência, além da qualificação, na inclusão e manutenção no mercado de trabalho. Como não existe nenhuma forma direta de mensurar esta dimensão do capital humano, consideramos como proxy do que denominamos “experiência” a idade do indivíduo, mas propriamente construímos três faixas de idade no intervalo etário de estudo: Adolescente (14-17), Jovem (18-24) e Jovem adulto (25-29).

Para análise do estado ocupacional, dentro do arcabouço do modelo de Search & Matching, é conveniente trabalhar apenas com jovens dentro da força de trabalho, o que nos levou à seleção das seguintes variáveis ocupacionais:

- *VD4001*: Condição em relação a força de trabalho (na força de trabalho, 1; fora da força de trabalho, 2)
- *VD4002*: Condição de ocupação (ocupada, 1; desocupada, 2)

Assim, criamos uma variável indicadora, *estado_ocupacional*, que assume valor 1 para as observações em que se está dentro da força de trabalho e ocupado, ou 0, para aquelas em que se está desocupado, mas ainda assim está dentro da força de trabalho. Ou seja, observar apenas jovens que estão ativamente no mercado de trabalho, seja trabalhando, seja buscando emprego.

Concluída esta etapa, de definição do conjunto de dados sobre os quais queremos proceder nossas análises dos impactos do choque econômico desencadeado pela pandemia de COVID-19 sobre a transição dos jovens no mercado de trabalho, passamos para a estimação dessas transições.

3. Construção da Matriz de Markov

Para a confecção das matrizes de transição a lá Markov, principiamos o emparelhamento de observações consecutivas do estado ocupacional de um mesmo indivíduo. Para isso, inicialmente ordenamos os dados por código de identificação individual, ano e trimestre, a fim de organizarmos observações do mesmo indivíduo uma abaixo da outra. Em seguida, agrupamos todas as observações pelo mesmo id, criamos variável *estado_t1* e armazenamos nela o estado ocupacional da observação posterior. Analiticamente isso significa,

$$E_{t1_t} = E_{t+1}$$

Depois, aplicamos um filtro para garantir que estejamos observando apenas registros consecutivos, ou seja, trimestre após trimestre, para que seja possível capturar as transições no espaço de um trimestre.

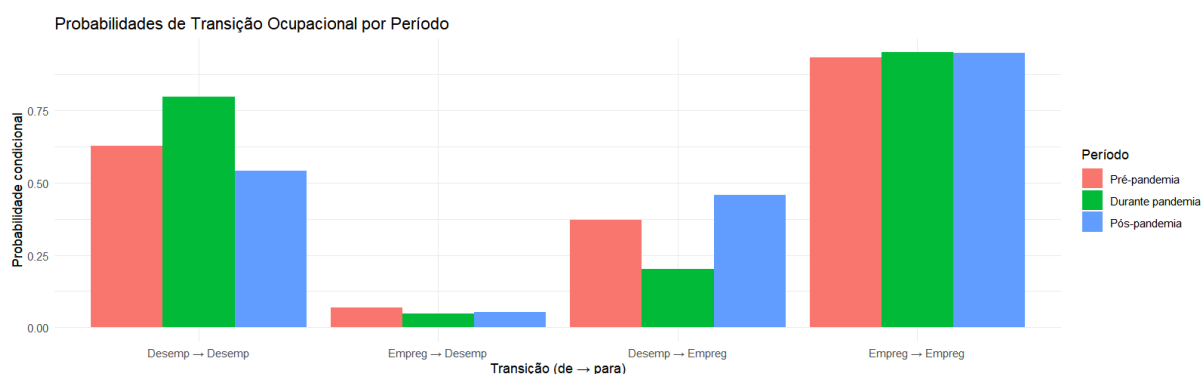
Com as informações contidas nestas duas variáveis, *estado_ocupacional* e *estado_t1*, conseguimos criar uma matriz que os unem como pares ordenados e contabilizam quantas vezes cada uma dessas combinações aparecem no conjunto de dados. Dividindo cada medida pelo total de observações, obtivemos as proporções, ou probabilidades condicionais de transição entre os estados ocupacionais 0 e 1, ou seja, entre o desemprego e emprego.

Como o que nos interessa são as matrizes de cada período de análise para comparação posterior, geramos uma função que estima matrizes de Markov para cada categoria da variável *período* especificada. Por exemplo, a seguinte tabela com matrizes para cada espaço temporal foi confeccionada:

Pré-pandemia			Durante pandemia		Pós-pandemia	
	Desemprego	Emprego	Desemprego	Emprego	Desemprego	Emprego
Desemprego	0,628	0,372	0,798	0,202	0,542	0,458
Emprego	0,068	0,932	0,049	0,951	0,052	0,948

A partir desses dados, já foram possíveis verificações de muitas afirmações da literatura consultada e adotadas como hipóteses de pesquisa, como a histórica dificuldade de inclusão juvenil no mercado de trabalho. Antes mesmo da pandemia, embora só esteja sendo observada as proporções de transição do ano de 2019, pode-se afirmar que, as chances de ingresso de um jovem no mercado de trabalho entre um trimestre e outro, controlada por todas as variáveis pessoais que possam impactar isso, são baixíssimas ($\approx 37,2\%$). Em média, a pandemia realmente teve um impacto negativo, reduzindo-as para cerca de 20,2%; o que pode ser explicado pela redução da atividade econômica, mas não pela rotatividade ou desemprego, já que a probabilidade de transição para o desemprego caiu, e a de permanência no emprego, aumentou (na verdade, nos três períodos, é muito alta). Já a explicação dos dados de retomada é um desafio a hipótese de trabalho, que supunha uma deterioração da participação dos jovens no mercado de trabalho; no entanto, as especulações não ficam de todas equivocadas, dado que a expansão dos serviços de delivery, especialmente por conta das restrições de aglomeração e trânsito durante a pandemia, tem uma fisionomia juvenil. Em outras palavras, o jovem encontrou nesse tipo de prestação de serviços um mercado laboral alternativo, dado a dificuldade de integração ao mercado formalizado.

Graficamente ficará melhor a visualização dessas chances de transição nos diferentes cenários estudados:



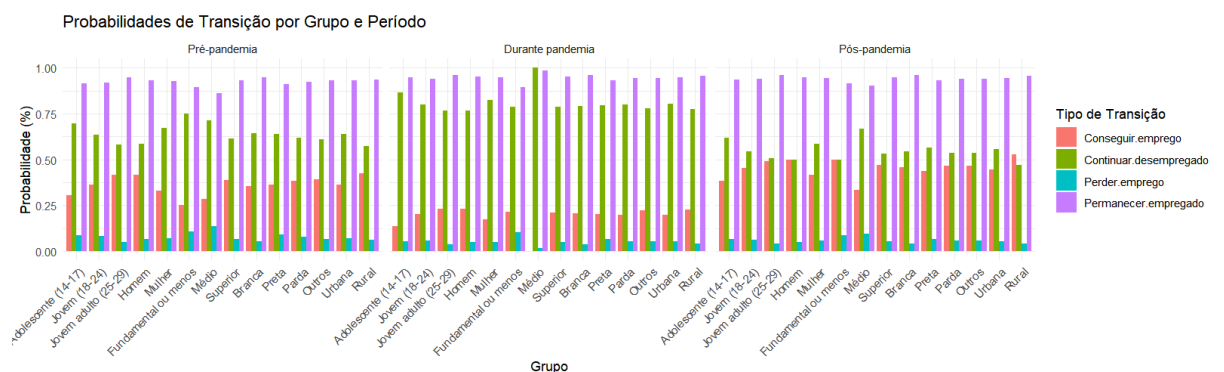
É visível que as chances de conseguir uma ocupação em um intervalo de três meses, para os jovens de 14 a 29 anos, é baixíssima, e, portanto, o tempo no desemprego é desalentador,

especialmente no período pandêmico. Mas a diferença entre as chances de ficar desempregado e de conseguir emprego se reduzem no pós-pandemia. Que mudanças na distribuição dos postos de trabalho, por exemplo, podem explicar esse fenômeno? Em pouco tempo as características do jovem que o tornam atrativo para o mercado de trabalho, e, portanto, dentro da visão do modelo DMP, aumentam a eficiência do matching, não devem ter mudado muito; mas, como se dá a distribuição dessas especificidades, e como ela impacta nas chances de o trabalhador conseguir uma ocupação nessa “nova economia”? Prosseguimos então com a geração de matrizes por grupo de características individuais.

3.1. Matrizes por Grupo e Período

Para geração dessas matrizes segmentadas tornamos explícitas as variáveis selecionadas no processo de filtragem, e criamos uma função a partir da qual possam ser geradas todas as matrizes de subgrupos.

O seguinte gráfico com as probabilidades de transição pode ser gerado:

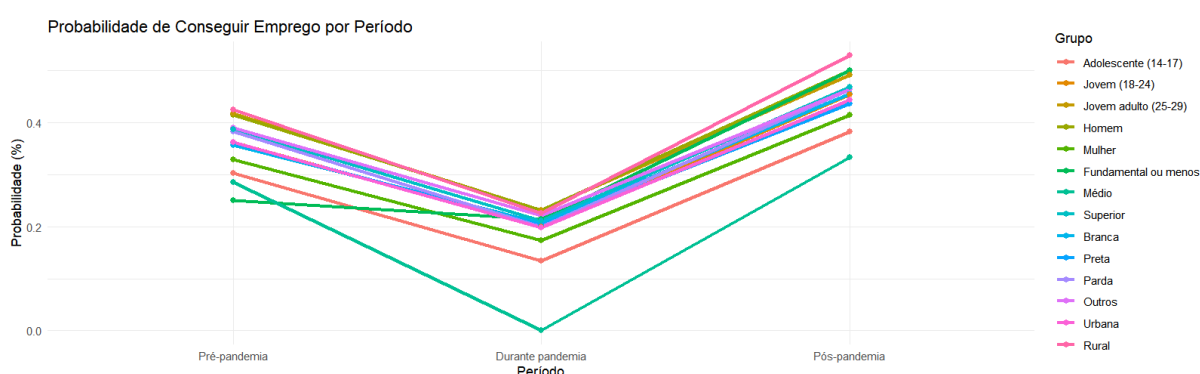


Quanto aos padrões de inclusão, segue-se o já observado na literatura, a pandemia apenas reduz as chances em escala, sem alterar as relações intragrupo. Quanto maior a idade, e na nossa abordagem, maior a experiência, maior as chances de conseguir um emprego; mulheres estão em desvantagem em relação aos homens; maior escolaridade acompanha maior integração, apesar de a diferença ser marginal entre quem tem Ensino Médio e quem tem apenas o Fundamental ou menos; jovens de áreas urbanas tem mais facilidade de transição para o emprego que jovens de regiões rurais, devido a um mercado de trabalho mais espesso.

No entanto, uma configuração nos parece interessante à investigação da pesquisa, pessoas de outras cores ou raças têm vantagens em relação às brancas, no processo de integração ao mercado de trabalho. Preliminarmente, pode-se afirmar que isto está associado a composição dos setores e postos de trabalho, a uma economia de empregos de baixa qualidade e amplo setor

informal, dado às desigualdades raciais consagradas na literatura de ciências sociais. Mas não se deve apegar a esta hipótese, pois os dados da pandemia podem esclarecer que não passa de um grande preconceito, e quem vai discordar dos números?

A pandemia, como já constatado, reduziu igualmente todas as probabilidades de inserção, mas significativamente achatou todas as chances associadas a categorias de raça e origem, elevou a adesão de quem tem fundamental ou menos, e significativamente zerou os de quem tem o ensino médio. Por outro lado, elevou todas as chances de ficar desempregado, e, portanto, o tempo no desemprego.



A retomada que se torna um objeto interessante da pesquisa, mais propriamente, as mudanças no durante pandemia, e o rearranjo da economia na recuperação, que explicam a melhora das chances de emprego, quando o que se vê na literatura é desmantelamento das cadeias de produção, crescente substituição da mão de obra humana por capital tecnológico e otimização de processos.

4. Modelagem Econométrica da Inclusão Juvenil no Mercado de Trabalho

Seguimos então para a modelagem econométrica dessas relações. Como o papel metodológico da pesquisa é aplicar o modelo Diamond-Mortensen-Pissarides ao estudo dos efeitos da pandemia, e, particularmente, o nosso interesse é mensurar essa reverberação no segmento juvenil da força de trabalho, empreenderemos uma estimação da eficiência do matching, que depende, além da disponibilidade de vagas e número de desempregados, de características dos trabalhadores que determinem sua eficiência e seu poder de barganha. Por isso, antes rodamos regressões Logit para as características, visando ver as que afetam a

probabilidade de transição do emprego com alguma significância estatística, inclusive analisando suas interações com o tempo.

Neste objetivo, inicialmente criamos a variável *transição_emprego* que foi nossa variável resposta nas especificações dos modelos. Ela assume valor 1 se a pessoa saiu do desemprego para o emprego, ou seja, se observações cuja variável *estado_ocupacional* é 0 têm como 1 o valor de *estado_t1*.

Após isso, como as especificações dos modelos foram tais que os coeficientes das variáveis explicativas fossem estimados em relação a um grupo base (homem, branca, fundamental ou menos, adolescente (14-17), urbano, pré-pandemia), criamos uma variável binária para mulher (que assume valor 1 se for; 0, caso contrário) e transformamos as outras e fator, tal que permita ao R criar dummies para cada grupo automaticamente.

A primeira especificação do modelo não considerou a interação com o tempo, e assumiu a seguinte forma:

$$P(Y_{it+1} = 1|X_{it}) = \text{logit}^{-1}(X)_{it}'$$

Onde:

- Y_{it+1} : variável dependente *transição_emprego* = 1 se ocorreu transição de desempregado para empregado.
- X_{it} : vetor de características individuais (sexo, raça, escolaridade, etc.) no tempo t.

Os coeficientes de cada variável assumem a forma de log-odds, de forma que se pode dizer que, o coeficiente da variável binária mulher é o efeito de ser mulher comparado a homem, que o coeficiente de *periodoDurante pandemia* é o efeito da pandemia comparado ao pré-pandemia, e assim por diante.

Em código R, a regressão tem a seguinte fórmula:

```
1. modelo_logit <- glm(transicao_emprego ~ mulher + raca + escolaridade + localidade +  
2. experiencia + periodo, data = microdados_transicao, family = binomial)
```

Para obter os coeficientes Odds-ratios (o quanto a elasticidade de uma variável representa da elasticidade da variável base), aplicou-se o exponencial aos coeficientes em log-odds obtidos do sumário da regressão. Ademais, considerou-se significativo apenas estimativas a um grau de 0.05 de significância. Abaixo, o sumário obtido:

Sumário do Modelo Logit Inicial

Variável Explicativa	Coef. Estimado	Odd-ratio	Variação odd	Erro padrão	Significância
Intercepto	-3,241771	0,039094597	-96,09%	0,418956	1,01E-14
Mulher	0,142067	1,152653874	15,27%	0,014793	<2,2e-16
Raça Preta	0,414842	1,514131489	51,41%	0,025315	<2,2e-16
Raça Parda	0,352104	1,42205641	42,21%	0,016389	<2,2e-16
Raça Outras	0,324429	1,38324059	38,32%	0,074964	1,51E-05
Ens. Médio	-0,316711	0,728541273	-27,15%	0,717351	0,65885
Ens. Superior	0,432454	1,541034586	54,10%	0,415477	0,29794
Urbana	0,23352	1,263038088	26,30%	0,018342	<2,2e-16
Jovem (18-24)	-0,167972	0,845377505	-15,46%	0,053886	0,001826
Jovem adulto (25-29)	0,618994	1,857058901	85,71%	0,054413	<2,2e-16
Durante pandemia	-0,459	0,631915245	-36,81%	0,016709	<2,2e-16
Pós-pandemia	-0,189145	0,827666486	-17,23%	0,018754	2,05E-06
AIC	154742,1				
Pseudo-R²	0,0164				
N (observações)	380099				

Estes resultados nos permitem fazer algumas inferências:

- Raças não-brancas tem consistentemente mais chances de conseguir emprego.
- Mulheres, que historicamente tinham menos chances, estão com 15% a mais que os homens.
- O maior dinamismo urbano permite aos jovens dessas regiões uma vantagem de 26% sobre os das áreas rurais.
- Jovens adultos (25-29) têm grande dificuldade de inserção, e por outro lado, quanto mais novo menor a desvantagem em relação aos adolescentes.
- Período da pandemia teve efeito fortemente negativo (-37%), com recuperação parcial no pós (-9%);
- Ter ensino superior ou ao menos o Médio, nesse período, não altera significativamente as chances de transição comparado ao fundamental.

Aqui se confirma, para o período em análise, que, as desigualdades raciais têm pouco efeito sobre o processo de inserção no mercado de trabalho. Como também, que, houve uma melhora significativa na presença das mulheres; o que pode ser parcialmente explicado pela expansão de setores em que elas compõem uma porção significativa da força de trabalho, como o setor de saúde e de cuidados.

No sentido contrário do observado na análise das matrizes de transição por experiência, os resultados da regressão mostram uma maior dificuldade dos mais velhos em aos adolescentes para ingressar no mercado de trabalho. No período pandêmico, especificamente, podemos supor que os mais velhos, devido ao maior risco de vida, foram preteridos aos mais novos, e isto resulte em um maior tempo de desemprego; ou, maior informalidade da economia, que

implique, agora seguindo o modelo de Search do McCall, em maior rejeição de ofertas pelos mais velhos, que são mais exigentes em termos de salário e estabilidade.

Claramente as áreas rurais foram as mais afetadas pela redução da atividade econômica, especialmente devido às medidas de restrição.

De fato, mais uma confirmação é feita do choque negativo da pandemia sobre a ocupação dos jovens no mercado de trabalho. E, de acordo com as estimativas, escolaridade não determinou maior ou menor grau de vulnerabilidade ao choque.

Ainda no esforço de entender o papel da heterogeneidade dos agentes no seu desempenho no mecanismo de matching, avançamos em uma especificação que considere as interações das características com os períodos. Ou seja, caberá a pergunta se, estatisticamente, ser mais jovem, e, portanto, ter menos experiência, ou ter um nível de escolaridade menor e - se considerarmos ter alguma relação causal - ser menos produtivo reduz as chances de admissão, em um cenário de redução da atividade econômica e muitas incertezas.

Assim adotamos um modelo DIF-IN-DIF, ou seja, que isole Diferenças em Diferenças com Heterogeneidade. Sua especificação foi a seguinte:

```
1. modelo_3fases <- glm(transicao_emprego ~ pandemia_fase * (mulher + raca + escolaridade +
2. localidade + experiencia), data = microdados_transicao, family = binomial)
```

Na próxima página, o sumário da regressão:

Modelo Dif-in-Dif					
Variável Explicativa	Coef. Estimado	Odd-ratio	Variação odd	Erro padrão	Significância
Intercepto	-4,194151	0,0150835	-98,49%	-4,141	3,45E-05
Mulher	0,156468	1,1693738	16,94%	6,8036	1,02E-11
Raça Preta	0,488409	1,6297217	62,97%	12,256	<2,2e-16
Raça Parda	0,440727	1,5538369	55,38%	17,100	<2,2e-16
Raça Outras	0,241820	1,2735656	27,36%	1,9302	5,36E-02
Ens. Médio	1,282137	3,6043364	260,43%	1,0296	3,03E-01
Ens. Superior	1,298107	3,6623583	266,24%	1,2861	0,198422

Urbana	0,256428	1,2923066	29,23%	9,0577	<2,2e-16
Jovem (18-24)	-0,158676	0,8532723	-14,67%	-1,989	0,303195
Jovem adulto (25-29)	-0,625733	0,5348691	-46,51%	-7,756	0,198422
Durante pandemia	0,772659	0,5216063	-47,84%	0,658	<2,2e-16
Pós-pandemia	1,360137	1,0361756	3,62%	1,0866	0,046691
Mulher x Durante pandemia	-0,650842	0,8856247	-11,44%	-0,914	8,71E-15
Mulher x Pós-pandemia	0,035536	0,8837556	-11,62%	0,9316	5,11E-01
Raça Preta x Durante pandemia	-0,121462	0,8856247	-11,44%	-2,069	0,277216
Raça Preta x Pós-pandemia	-0,123574	0,8837556	-11,62%	-1,919	0,055524
Raça Parda x Durante pandemia	-0,168657	0,8447984	-15,52%	-4,483	0,351535
Raça Parda x Pós-pandemia	-0,121130	0,8859185	-11,41%	-2,856	0,038474
Raça Outras x Durante pandemia	0,088755	1,0928129	9,28%	0,4953	0,054893
Raça Outras x Pós-pandemia	0,177558	1,1942979	19,43%	0,9563	7,34E-06
Ens. Médio x Durante pandemia	-10,35933	3,16955E-5	-100,00%	-0,277	0,004278
Ens. Médio x Pós-pandemia	-1,659437	0,1902460	-80,98%	-0,938	0,620398
Ens. Superior x Durante pandemia	-1,082614	0,3387087	-66,13%	-0,927	0,338896
Ens. Superior x Pós-pandemia	-1,308413	0,2702485	-72,98%	-1,051	0,781539
Urbana x Durante pandemia	-0,025127	0,9751857	-2,48%	-0,595	0,348191
Urbana x Pós-pandemia	-0,060047	0,9417202	-5,83%	-1,272	0,353821
Jovem (18-24) x Durante pandemia	-0,020860	0,9793552	-2,06%	-0,165	0,293078
Jovem (18-24) x Pós-pandemia	-0,007340	0,9926862	-0,73%	-0,053	0,5514992
Jovem adulto (19-24) x Durante pandemia	0,0574106	1,0590905	5,91%	0,4518	0,203091
Jovem adulto (19-24) x Pós-pandemia	-0,056524	0,9450437	-5,50%	-0,41	0,868419
AIC	169160,2				
Pseudo-R² (McFadden)	0,0167				
N (observações)	380099				

Os resultados desta estimação corroboram os das anteriores. As generalizações possíveis estão organizadas na tabela abaixo:

Grupo	Durante Pandemia	Pós-pandemia
Mulheres	Penalizadas marginalmente (~6%)	Recuperadas (sem efeito no pós)
Pardos	Penalizados significativamente (~15%)	Penalização persiste (~11%)
Pretos	Penalizados (~11%)	Recuperação no pós

Outras Raças	Sem efeito claro	Sem efeito claro
Escolaridade	Sem interação significativa	—
Capital	Benefício mantido, sem mudança	—
Jovens (18–29)	Desvantagem consistente, mas sem mudança ao longo dos períodos	

Podemos concluir então que a pandemia penalizou especialmente os jovens pardos e, em menor grau, pretos e mulheres; e que, apenas as mulheres se recuperaram plenamente no pós-pandemia.

5. Estimação da função de Matching e de sua eficiência entre os jovens

Para estimar o impacto do choque pandêmico sobre a eficiência com que os jovens trabalhadores brasileiros conseguem encontrar uma vaga de emprego, aplicamos empiricamente o modelo de matching, com base no arcabouço teórico de Diamond, Mortensen e Pissarides (DMP).

A partir da comparação entre a taxa observada de transição do desemprego para o emprego, estimada nos processos anteriores de confecção das Matrizes de Transição, e a tightness do mercado de trabalho juvenil, identificamos se há ineficiências estruturais no matching deste grupo, e como as ineficiências evoluíram ao longo do tempo, especialmente seu comportamento durante a pandemia.

5.1. Modelo Teórico: Função de Matching Cobb-Douglas

Para determinação da função de matching assumimos uma forma de uma função de produção que relaciona o número de novas admissões M_t com o estoque de desempregados U_t e de vagas V_t :

$$M_t = m \cdot U_t^\alpha \cdot V_t^{1-\alpha}$$

Onde:

- M_t : número de admissões (ou matches) no trimestre t
- U_t : número de desempregados
- V_t : número de vagas
- m : eficiência tecnológica da função de matching
- $\alpha \in (0,1)$: elasticidade em relação a U_t

Para a estimação empírica, dividimos os lados por U_t , obtendo a taxa de encontro f_t , ou seja, a proporção de admissões por número de desempregados:

$$f_t = \frac{M_t}{U_t} = m \cdot \left(\frac{V_t}{U_t}\right)^{1-\alpha}$$

Tomando logarítimo:

$$\log(f_t) = \log(m) + (1 - \alpha)\log(\theta_t)$$

Onde:

- f_t : taxa de transição do desemprego para o emprego
- $\theta = V_t/U_t$: tightness (pressão no mercado de trabalho)

Esta conseguimos facilmente estimar via regressão linear:

$$\log(f_t) = \beta_0 + \beta_1 \log(\theta_t) + \epsilon_t$$

Sendo:

- $\beta_1 = 1 - \alpha \Rightarrow$ permite recuperar α
- $\beta_0 = \log(m)$

5.2. Estratégia de estimação empírica

Diferentemente do caminho utilizado para estimar as matrizes de transição e as contribuições estatísticas das características individuais para as chances de transição, consideramos um banco de dados mais amplo, para todas as idades, e para um período maior (2012-2023), de forma a ter pontos de dados suficientes para ter um estimador minimamente significativo.

A partir disso, geramos um painel trimestral agregado, com as variáveis indicadoras:

- U_t : total de indivíduos desempregados
- E_t : total de indivíduos ocupados
- M_t : número de indivíduos que estavam desempregados no trimestre t e foram ocupados no trimestre $t + 1$
- $f_t = M_t/U_t$: taxa de encontro

Também, as observações foram emparelhadas a fim de identificar o estado ocupacional de cada indivíduo em dois trimestres consecutivos.

Como os número de vagas abertas em cada período são disponibilizados diretamente, utilizamos como proxy o número de admissões formais mensais do antigo e novo CAGED, e agregamos por trimestre. Para o modelo juvenil, utilizamos dados de admissões por faixa etária disponibilizados pelo mesmo sistema. Assim pudemos estimar thightness separadamente para toda a população e apenas para os jovens.

Dado o parâmetro $\hat{\alpha}$ estimado com o modelo geral, estimamos a eficiência relativa de matching dos jovens como:

$$Ef f_t^{jovem} = \frac{f_t^{jovem}}{(\theta_t^{jovem})^{1-\hat{\alpha}}}$$

Onde:

- f_t^{jovem} : taxa de encontro dos jovens
- θ_t^{jovem} : tightness do mercado jovem
- $\hat{\alpha}$: elasticidade estimada do matching geral

Intuitivamente, quanto maior as barreiras ou fricções no mercado jovem, menor a eficiência de matching, e, portanto, esse indicador assumirá um valor menor que 1, apontando distanciamento da eficiência padrão esperada pelo modelo.

5.3. Eficiência de Matching dos Jovens estimada

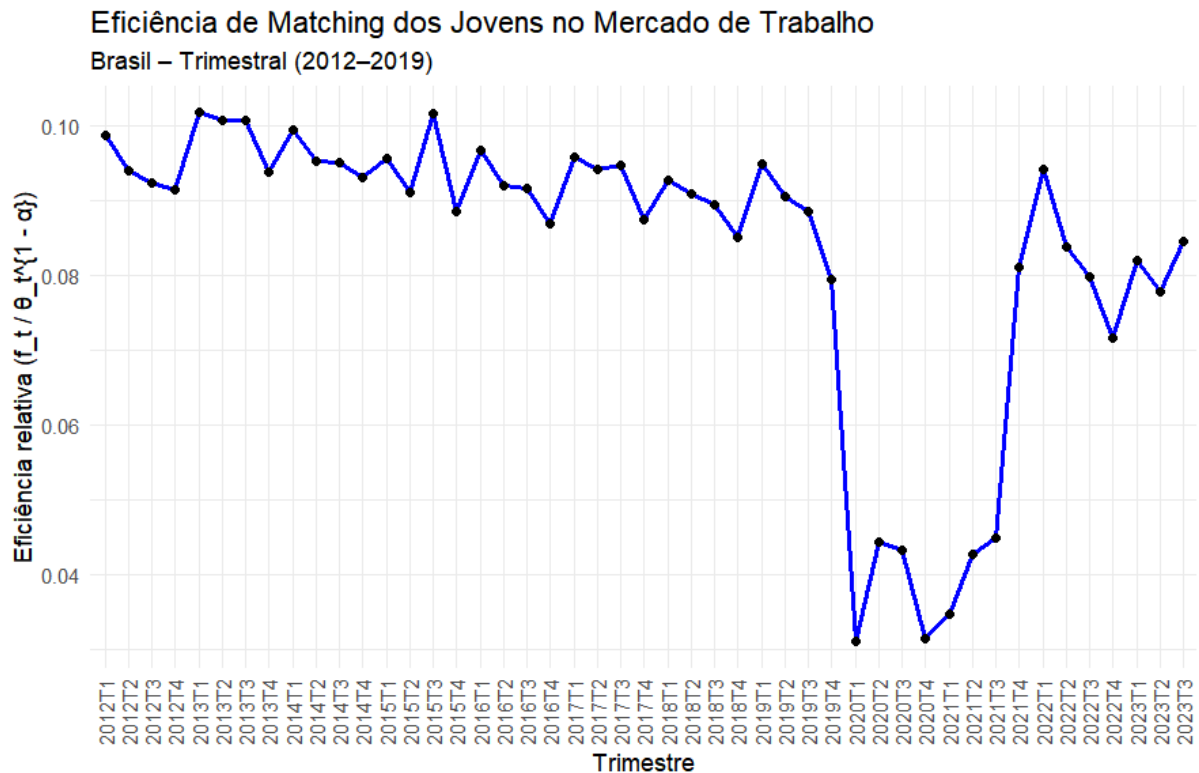
A partir desse modelo conseguimos obter resultados muito expressivos e relevantes, chegando ao seguinte formato logarítmico:

$$\log(f_t) = -2.4398 + 0.2650 \cdot \log(\theta_t) + \epsilon_t$$

A relação entre a tightness e a taxa de encontro é estatisticamente significativa ao nível de 5% ($p < 0.05$), e nos permitiu extrair o parâmetro estrutural da função de matching ($\hat{\alpha} = 1 - \hat{\beta}_1 = 0.735$), que representa a elasticidade do matching em relação ao número de desempregados.

Com base nesse parâmetro, estimamos a eficiência de matching entre os jovens, a razão entre a taxa de transição observada e a prevista pelo modelo. Assim pudemos observar a

capacidade do mercado de trabalho de integrar jovens ao emprego, dado o grau de escassez relativa de vagas. E, construímos a seguinte série temporal:



A análise dessa série temporal nos revelou padrões relevantes:

- Entre 2012 e 2019, a eficiência manteve-se relativamente estável, com leves oscilações em torno de 0,09 a 0,10, mesmo durante a recessão de 2015-2016.
- A partir do 1º trimestre de 2020, observou-se uma queda abrupta da eficiência, que atingiu seu ponto mais baixo no 2º trimestre de 2020, em meio ao choque da pandemia de COVID-19.
- Embora tenha havido uma recuperação parcial nos trimestres seguintes, a eficiência não retornou aos níveis anteriores, permanecendo em patamar inferior até 2023.

Esse resultado revela que, embora a tightness do mercado de trabalho juvenil tenha efetivamente caído durante a pandemia — em razão da redução de vagas e do aumento do desemprego —, a taxa de contratação dos jovens caiu mais do que o esperado, dado esse nível de tightness. Isso se expressa diretamente na queda da eficiência, sugerindo a presença de fricções adicionais no mercado de trabalho juvenil durante a crise sanitária.

Tais fricções podem estar associadas à natureza dos vínculos empregatícios interrompidos (setores de alta rotatividade), à dificuldade de buscar ativamente trabalho em meio a restrições sanitárias, ou ainda a mudanças no perfil da demanda por trabalho (retração de setores intensivos em mão de obra jovem).

Também prosseguimos, dentro da disponibilidade de dados, gerando estimativas de eficiência de matching para os jovens por sexo e escolaridade. No entanto, a visualização da série temporal por escolaridade não ficou adequadamente apresentável, por pontos críticos nos dados.

