

O Papel das Redes Sociais na Condição do Desempregado: Uma Análise Usando a Teoria de *Matching*

Francis Carlo Petterini

Professor do Programa de Pós-Graduação em Economia da Universidade Federal de Santa Catarina (PPGECO/UFSC), Brasil

Resumo

O artigo investiga se o tamanho das redes sociais pode afetar a taxa de desemprego da economia. Para tanto, com base em um modelo teórico de *matching*, propõe-se uma técnica de simulação dos efeitos das redes na empregabilidade dos trabalhadores. A proposta é aplicada em uma base de dados de Portugal, observando-se evidências que de que existe um tamanho ótimo de rede que minimiza a taxa de desemprego.

Palavras-chave: Redes Sociais, Desemprego, *Matching*

Classificação JEL: J64, C63, D01

Abstract

The paper investigates whether the size of social networks can affect the unemployment rate of the economy. Therefore, based on a theoretical model of matching, I propose a technique for simulating the effects of networks on the employability of workers. The proposal is applied to a database of Portugal, noting evidence that there is an optimal size of network that minimizes the unemployment rate.

* Recebido em setembro de 2011, aprovado em novembro de 2012. O autor agradece ao parecerista anônimo pelas contribuições.

E-mail address: f.petterini@ufsc.br

1. Introdução

Um considerável interesse tem emergido na literatura sobre como as variáveis econômicas se relacionam com as interações sociais. Isto é fundamental, em particular, para a economia do trabalho.

Do lado dos trabalhadores, existem diferentes níveis de produtividades (associadas as diferentes níveis de educação, habilidades etc.), diferentes intensidades e formas de busca por emprego (via agências, anúncios de jornal etc.), diferentes salários de reserva (decorrentes das preferências em relação ao trabalho e lazer), diferentes hábitos culturais etc. Do lado das firmas, existem diferentes intensidades do uso de capital e trabalho, diferentes técnicas de seleção de pessoal etc. E de ambos os lados estes aspectos são notadamente influenciados pelo ambiente social onde se encontram trabalhador e firma.

Uma proeminente forma de se tratar essas idiosincrasias no mercado de trabalho surgiu com a *teoria de matching* na década de 1970. Tal forma de análise nasce da adaptação dos processos estatísticos de *urna-bola* feitos nos trabalhos seminais de Robert Hall, Peter Diamond, Dale Mortensen e Christopher Pissarides (ver Petrongolo e Pissarides 2001). A teoria diz que trabalhadores desempregados podem ser vistos como bolas e vagas de emprego como urnas. Em cada momento do tempo existe um processo que lança as bolas em direção as urnas. Ao final do processo, algumas bolas estarão pareadas (*matching*) com algumas urnas, mas sempre existirão bolas perdidas e urnas vazias.

O artigo analisa o processo de *matching* em razão do tamanho de uma rede social (um conjunto de trabalhadores que pode indicar um posto de trabalho vago a um desempregado). Neste sentido, Calvó-Armengol (2004) e Calvó-Armengol e Zenou (2005) apontaram a existência teórica de um tamanho ótimo de rede no processo, capaz de minimizar a taxa de desemprego de uma economia.

Esse resultado é difícil de se testar empiricamente, uma vez que é difícil de se observar o tamanho das redes sociais. Mas o artigo propõe um modelo *mixed-logit* (Train 2009) para simular as escolhas das firmas de como preencher um posto de trabalho com determinados tipos de trabalhadores. Onde os desempregados podem (ou não) fazer uso da sua rede social para melhorar a empregabilidade. A proposta é aplicada usando uma base de dados para Portugal, e os resultados estimados corroboram com a tese que existe um tamanho ótimo de rede social para minimizar as taxas de desemprego.

Além desta Introdução, a Seção 2 contextualiza o artigo. A Seção 3 apresenta a estratégia empírica proposta. A Seção 4 descreve a base de dados. A Seção 5 apresenta os resultados do *mixed-logit* proposto e das simulações conduzidas. Por fim, a Seção 6 aponta as Considerações Finais.

2. Contextualização

Entende-se por *rede social* um agrupamento de pessoas com interesses em comum. A rede é responsável pelo compartilhamento de informações. No artigo, a rede é definida como um conjunto de trabalhadores que pode indicar um posto de trabalho vago a um desempregado.

Com base em Ioannides e Loury (2004), a literatura que trata das redes sociais poderia dividida em três fases. A primeira começa com o trabalho de Rees (1966), documentando que muitos trabalhadores encontram seus empregos pela influência de amigos.

Com o trabalho de Rees, a literatura passou a analisar a busca por emprego em dois canais: o formal (ou convencional), que são as agências de emprego, os anúncios de jornal, a entrega direta de currículos etc.; e o informal (ou não convencional), que são as informações ou indicações sobre postos de trabalho recebidas de amigos e familiares.

Nesse sentido, parte da literatura focou em estabelecer fatos estilizados sobre o assunto. Os trabalhos de Corcoran et alii (1980), Datcher (1983), Bartlett e Miller (1985), Holzer (1988) e Topa (2001) apontam que três destes fatos podem ser estabelecidos para os Estados Unidos:

- 1) cerca de 1/4 dos trabalhadores usam sua rede social como principal meio de busca de emprego;
- 2) o uso e a eficiência da rede em aumentar as chances de *matching* podem variar substancialmente entre estratos da população; e,
- 3) trabalhadores que usam sua rede podem ter maiores chances de *matching*, mas isto não implica em salários diferenciados.

A segunda fase se caracterizaria com a publicação dos primeiros modelos teóricos que tratam da influência das redes sociais no mercado de trabalho. Esta fase é marcada pelos trabalhos de Montgomery (1991, 1994), com resultados indicando que quanto maior a rede, menor a chance de se estar desempregado.

Iniciando a terceira fase, os trabalhos de Calvó-Armengol (2004) e Calvó-Armengol e Zenou (2005), doravante denominados por modelo CAZ, indicaram que a relação entre rede social e mercado de trabalho é mais complexa e ambígua do que era postulado nos modelos influenciados por Montgomery.

O modelo CAZ é desenvolvido sobre uma estrutura de *matching* ao estilo Pissarides (2000). Para ser algebricamente tratável, os trabalhadores são considerados homogêneos (mesma educação, habilidade etc.) e existem duas suposições de base:

- 1) um trabalhador empregado que conhece um posto de trabalho vago só detém esta informação durante o período vigente (ou, de forma alternativa, se uma vaga não for preenchida no período vigente ela é destruída); e,
- 2) um trabalhador recentemente empregado passa por um período probatório, quando sua produtividade é (normalizada em) zero.¹

¹ É uma fase na qual o ele apenas aprende a cultura da firma. O trabalhador passa a ser produtivo só

Juntas, estas duas suposições implicam que um trabalhador empregado que ouve falar de um posto desocupado não terá receio de transmitir a informação. Pois, como as vagas são válidas por apenas um período e existem diferenças de salário decorrentes das diferenças de produtividade, nenhum trabalhador empregado tem incentivo para esconder informação.

Em cada ponto do tempo existem n trabalhadores, que podem estar empregados ou desempregados. E, no período t , a taxa de desemprego é u_t .

As firmas anunciam as vagas de forma que todos os trabalhadores têm igual acesso a informação. Então, cada trabalhador pode isoladamente se defrontar com um determinado posto de trabalho desocupado com probabilidade $v_t = V_t/n$ (onde V_t é o número de postos vagos no momento).

Mas o trabalhador pode estar inserido em uma rede social de tamanho s . Então, a cada período cada trabalhador está potencialmente em contato com s dos n trabalhadores. Calvó-Armengol (2004) mostra que a probabilidade do desempregado encontrar uma vaga pela indicação da rede é:

$$p_t = 1 - \left(1 - v_t(1 - u_{t-1}) \frac{1 - (1 - u_{t-1})^s}{su_{t-1}} \right)^s \quad (1)$$

O resultado acima é a chave do modelo CAZ, pois para cada par (u, v) existe um único máximo global para p_t em função do tamanho da rede social. Isto ocorre porque o aumento de s gera dois efeitos em p_t . O primeiro é que ao aumentar a rede o desempregado aumenta a chance de encontrar um trabalhador empregado capaz de lhe indicar uma vaga. O segundo é que ao aumentar a rede o desempregado também aumenta a chance de que a vaga seja indicada para outro desempregado. E estes efeitos crescem em proporções diferentes em relação a s .

Esse é um resultado intrigante. Pois, em outras palavras, sugere que se o desempregado possui poucas conexões sociais, na medida em que ele aumenta estas conexões, também aumentam as chances dele encontrar emprego. Mas como a rede também conta com outros desempregados, se ela aumenta demasiadamente a informação do trabalhador empregado (que indica uma vaga) passa a ser compartilhada por um grupo maior de trabalhadores desempregados. Isto gera uma espécie de externalidade negativa de estar em uma rede muito grande, pois a notícia de uma mesma vaga chega a muitos desempregados. E a chance de cada um individualmente se empregar se reduz.

Quanto a estrutura de *matching*, a probabilidade de um trabalhador desempregado ser contratado no período t é a probabilidade de: ser contratado da forma convencional; ou não achar uma vaga da forma convencional e achar via rede social. Sendo computada por: $v_t + (1 - v_t)p_t$. Como existem u_{t-1} trabalhadores desempregados no início do período, a probabilidade de que ocorra um *matching* nesta economia no decorrer do período é dada por: $m_t = u_{t-1}(v_t + (1 - v_t)p_t)$. E

após o período de provação, e mantém sua produtividade constante até que se desligue do emprego. E quando ele achar um novo trabalho passará por um novo período probatório.

de forma análoga ao que acontece com p_t , para cada par (u, v) , m_t também possui um único máximo global em função do tamanho da rede.

Como consequência dos resultados anteriores, Calvó-Armengol e Zenou (2005) também demonstram que o mesmo número de conexões sociais que maximiza p_t , *ceteris paribus*, é aquele que minimiza a taxa de desemprego de toda a economia. Mais ainda, mostraram que em equilíbrio de estado estacionário a taxa de desemprego é uma função do tamanho da rede social, $u(s)$, tal que: $u(0) < u(s) < u(n)$ ou $u(n) < u(s) < u(0)$, $\forall s \neq \{0, n\}$.

Esse último resultado é complicado de se traduzir em termos de política de emprego, uma vez que a política teria de orientar o número de conexões sociais dos trabalhadores. Mas pode explicar parte de uma inflexibilidade em se reduzir desemprego. Pois indica que duas economias iguais em todos os fatores, exceto pela forma das redes sociais, podem ter taxas de desemprego diferentes.

3. Estratégia Empírica

O modelo CAZ é difícil de se testar empiricamente, pois é difícil observar o tamanho das redes sociais. Mas Wahba e Zenou (2005) propuseram uma estratégia, definindo duas variáveis binárias para estimar um modelo *logit* simples: S se o trabalhador usa a rede para procurar o emprego, e E se o trabalhador consegue emprego. Em um exercício empírico usaram uma base de dados do Egito, onde cerca de 1/3 dos empregados afirmaram ter encontrado a vaga devido a indicações de amigos ou familiares (ou seja, usaram sua rede). Pela proposta, a probabilidade do trabalhador estar empregado afirmando que usou a rede, $\Pr(E = 1, S = 1)$, seria uma *proxy* para p_t – o resultado da equação (1).

A matriz de controles de Wahba e Zenou possuía um vetor que apontava a densidade populacional do local onde estava o trabalhador, definida por \tilde{s} . Supostamente, a densidade seria uma *proxy* para s . E então estimaram o modelo *logit*: $\Pr(E = 1, S = 1) = \exp(\beta_0 + \beta_1\tilde{s} + \beta_2\tilde{s}^2)/(1 + \exp(\beta_0 + \beta_1\tilde{s} + \beta_2\tilde{s}^2))$. Os resultados apontaram $\beta_1 > 0$ e $\beta_2 < 0$, corroborando com o previsto no modelo CAZ. Pois observa-se um máximo para $\Pr(E = 1, S = 1)$ em relação a \tilde{s} .

A estratégia empírica deste artigo é uma alternativa da análise de Wahba e Zenou (2005), que não usa uma *proxy* para s . Propõe-se simular cenários de escolhas das firmas onde os trabalhadores são (ou não) indicados pela rede. E mais do que estudar a probabilidade do trabalhador se empregar usando a rede, a variável de impacto analisada será a taxa de desemprego da economia em função de variações nas redes sociais. E para tanto, se recorrerá a um sistema de numeração posicional (ver Scheinerman 2005).

Como ilustração, imagine que o pesquisador observa uma a matriz X de covariadas. Nas linhas estão as informações de cada indivíduo e nas colunas está uma sequência de K *dummies* que indicam características do trabalhador. Por exemplo, tome $K = 3$ da seguinte forma: a primeira coluna indica o gênero (1 para homem e 0 para mulher), a segunda a faixa etária (1 para menos de 30 anos e 0

caso contrário) e a terceira a forma com que o emprego foi encontrado (1 via rede social e 0 caso contrário).

Neste caso existem $2^3 = 8$ possibilidades de tipos de trabalhadores observados através das linhas de X . Onde: uma posição binária 0|1|0 significa mulher com menos de 30 anos que não encontrou emprego via rede social; e uma posição binária 1|0|1 significa homem de mais de 30 anos empregado usando a rede social.

Assim, se fragmentará o mercado de trabalho por K características, em até 2^K tipos, indexados por $j = 0, 1, \dots, 2^K - 1$. A regra de indexação é dada por $j = \sum_{k=1}^K 2^{k-1} d_k$, onde d_k é uma *dummy* indicando a presença das características $k = 1, \dots, K$. E $j = 0$ é chamado de *tipo base*, quando todas as *dummies* são zero. Como é possível que não se observe todos os tipos em um dado momento, define-se por J o conjunto dos tipos observados.

O sistema de numeração posicional pode ser facilmente aplicado usando os dados de pesquisas de emprego e desemprego. Mas para existir *matchings* a serem estudados, além dos trabalhadores é preciso que se tenham vagas de emprego.

Para resolver o lado dos empregos se usará o modelo *mixed-logit*. Seu papel é simular os postos de trabalho nos seguintes termos: existem T postos (ocupados ou não); T não precisa ser conhecido pelo pesquisador, tampouco cada posto precisa ser identificado; o pesquisador não sabe por qual perfil j cada um dos $\tau = 1, \dots, T$ postos está sendo ocupado (ou se não está sendo ocupado), mas sabe que as firmas escolhem o melhor para si.

Suponha então que cada firma tem uma função de escolha $E_{j\tau}$, que conecta o tipo de trabalhador j com o posto de trabalho τ . Defina esta função por:

$$E_{j\tau} = \underbrace{\beta_{0\tau} + \sum_{k=1}^K \beta_{k\tau} d_{kj}}_{\Omega_{j\tau}} + \xi_j + \varepsilon_{j\tau} \quad (2)$$

Onde: $\beta_{0\tau}$ é uma constante que representa a ponderação do tipo base para o emprego τ ; $\beta_{k\tau}$ é o peso dado a característica k para o emprego τ ; ξ_j é um índice que agrega outras características do tipo que não puderam ser captadas pelo sistema de numeração posicional; e $\varepsilon_{j\tau}$ é um termo que indica a subjetividade da escolha da firma, que não é observada pelo pesquisador.

A firma pode escolher não ocupar a vaga. Neste caso define-se a escolha pelo subscrito u , tal que: $E_{u\tau} = \Omega_{u\tau} + \varepsilon_{u\tau}$. Como argumentado em McFadden e Train (2000) e McFadden (2001), normalizando $\Omega_{u\tau} = 0$ pode-se assumir, sem perda de generalidade, que o termo não-observável $\varepsilon_{j\tau}$ segue uma distribuição de probabilidade Gumbel.² E como demonstrado em Train (2009) para o caso geral das funções de escolha, conclui-se que a probabilidade de se encontrar um trabalhador do tipo j conectado ao posto de trabalho τ é dada por:

$$q_{j\tau} = \frac{\exp(\beta_{0\tau} + \sum_{k=1}^K \beta_{k\tau} d_{kj} + \xi_j)}{1 + \sum_{r \in J} \exp(\beta_{0\tau} + \sum_{k=1}^K \beta_{k\tau} d_{kr} + \xi_r)} \quad (3)$$

² Ou Valor Extremo Tipo I, com parâmetro de localização igual a zero e parâmetro de escala igual a unidade.

Considere então que $\bar{\beta}_0$ e $\bar{\beta}_k$ são, respectivamente, a constante média e o peso médio da característica k nas escolhas. Se σ_0 e σ_k são os respectivos desvios padrão destes pesos, então seguinte a normalização pode ser feita: $z_{0\tau} = (\beta_{0\tau} - \bar{\beta}_0)/\sigma_0$ e $z_{k\tau} = (\beta_{k\tau} - \bar{\beta}_k)/\sigma_k$. Estes termos seguirão distribuição normal padrão.

Considere z como um vetor de dimensão $T(K + 1)$, contendo os termos normalizados, e $F(z)$ como sua distribuição cumulativa de probabilidade conjunta. Então a probabilidade de que um trabalhador esteja empregado e seja tipo j será dada pela seguinte esperança matemática aplicada em (3):

$$q_j = \int_z \frac{\exp((\bar{\beta}_0 + \sum_{k=1}^K \bar{\beta}_k d_{kj} + \xi_j) + (\sigma_0 z_{0\tau} + \sum_{k=1}^K \sigma_k d_{kj} z_{k\tau}))}{1 + \sum_{r \in J} \exp((\bar{\beta}_0 + \sum_{k=1}^K \bar{\beta}_k d_{kr} + \xi_r) + (\sigma_0 z_{c\tau} + \sum_{k=1}^K \sigma_k d_{kr} z_{k\tau}))} dF(z) \tag{4}$$

Com base no resultado (4) e na característica de excludência do sistema de numeração posicional, a probabilidade de que algum trabalhador esteja empregado é dada por $\sum_j q_j = 1 - u$, ou seja, a taxa de emprego observada. Então, o lado esquerdo de (4) equivale $q_j = e_j/n$, onde e_j é o número de trabalhadores do tipo j que estão empregados e n é o número total de trabalhadores. Logo o lado esquerdo de (4) pode ser mensurado usando-se as bases de dados de pesquisas de emprego e desemprego. E como o lado direito da igualdade é uma esperança matemática, o resultado pode ser rescrito por:

$$\frac{e_j}{n} = \frac{1}{T} \sum_{\tau=1}^T \frac{\exp((\bar{\beta}_0 + \sum_{k=1}^K \bar{\beta}_k d_{kj} + \xi_j) + (\sigma_0 z_{c\tau} + \sum_{k=1}^K \sigma_k d_{kj} z_{k\tau}))}{1 + \sum_{r \in J} \exp((\bar{\beta}_0 + \sum_{k=1}^K \bar{\beta}_k d_{kr} + \xi_r) + (\sigma_0 z_{c\tau} + \sum_{k=1}^K \sigma_k d_{kr} z_{k\tau}))} \tag{5}$$

É claro, também há uma conexão entre q_j e a taxa de *matching*. Pois se o trabalhador está empregado no período t , ele incorreu em *matching* ou já estava empregado no período $t - 1$. Como é de praxe nestes modelos, assume-se uma taxa de destruição exógena dos postos de trabalho dada por $\lambda \in (0, 1)$. Deduzindo-se que: $1 - u_t = m_{t-1} + (1 - \lambda)(1 - u_{t-1})$. E replicando este resultado em períodos subsequentes, a probabilidade de um determinado trabalhador estar empregado pode ser escrita por: $1 - u_t = \sum_{i=0}^{\infty} (1 - \lambda)^i m_{t-1-i}$.

Como em Blanchard et alii (1989), considere a dinâmica de *matching* dada por $m_t = \bar{m} + \zeta_t$. Onde \bar{m} é uma constante que representa a probabilidade média de *matching* da economia ao longo do tempo e ζ_t é um ruído branco. Assim, conclui-se que a probabilidade média de um determinado trabalhador estar empregado pode ser escrita por: $1 - \mathbb{E}(u) = \bar{m}(1 + 1/\lambda)$.

Esse resultado aponta que a taxa média de emprego é a taxa média de *matching* amplificada por um fator inverso a taxa de destruição dos postos de trabalho.³ E

³ Por exemplo, usando os valores de Blanchard e Portugal (2001), têm-se como fatos estilizados para a década de 1990 um λ de 0,039 para Portugal e de 0,073 para os Estados Unidos. Assim, a taxa média de emprego em Portugal seria 26,6 vezes maior que a sua taxa média de *matching*, e nos Estados Unidos seria 14,7 vezes.

observando-se os resultados anteriores, é possível escrever uma conexão entre q_j e a taxa de *matching* dada por $\Sigma_j q_j = \bar{m}(1 + 1/\lambda)$.

Em princípio, o pesquisador desconhece o total de postos de trabalho da economia (T) e suas idiossincrasias de escolha (o vetor z). Mas a forma funcional (5) pode ser estimada através de simulações como descrito em Berry e Pakes (2007) e Train (2009), tornando-a operacional para um exercício empírico.⁴

Para acontecer os *matchings* é necessário existir desempregados (na proporção u) e vagas (na proporção v , seguindo a definição da seção anterior). Suponha então que o pesquisador estimou (5). Então, durante esta estimação foram criados T postos de trabalho simulados. E, em uma nova simulação, se pode transformar uma parte destes T postos em vagas.

Operacionalmente, criam-se as vagas $i = 1, \dots, V$ selecionando-as aleatoriamente entre as T usadas na estimação de (5). E cada vaga guarda as idiossincrasias do posto de trabalho que a originou (um vetor z_v de dimensão $K + 1$).

Esse é o aspecto fundamental da proposta, pois dependendo das idiossincrasias associados às vagas que surgem, mais (ou menos) trabalhadores desempregados podem incorrer em *matching* com estas vagas. Então o pesquisador pode verificar se aqueles que usam sua rede social para encontrar as vagas apresentam maiores chances de saírem do desemprego.

Para tanto desenham-se três cenários, e em todos define-se d_{Kj} como a *dummy* que indica se o tipo de trabalhador usa a rede social. O primeiro cenário é o referencial, quando se destrói $T - V$ postos de trabalho e computam-se as taxas de emprego resultantes para cada um dos J tipos. Esta taxa, denominada de q_j^v , é calculada através da equação (6), posta adiante – os acentos circunflexos indicam que os parâmetros foram estimados em um primeiro passo.

O segundo e o terceiro cenários são, respectivamente, aqueles onde nenhum e todo trabalhador faz uso da rede social. Logo, ou não se usa a ponderação para a rede, ou se usa com todo tipo de trabalhador. Cada cenário contém, respectivamente, J_s e J_c tipos de trabalhadores observados. Quando as taxas de emprego resultantes para cada tipo j , denominadas de q_j^s e q_j^c , são computadas através das equações (7) e (8), respectivamente.

⁴ Ilustrativamente, considere T um número fixo e um vetor z dado aleatoriamente. O lado esquerdo de (5) está fixo também, então pode-se buscar valores para os parâmetros da equação de forma a minimizar uma função dos termos ξ_j .

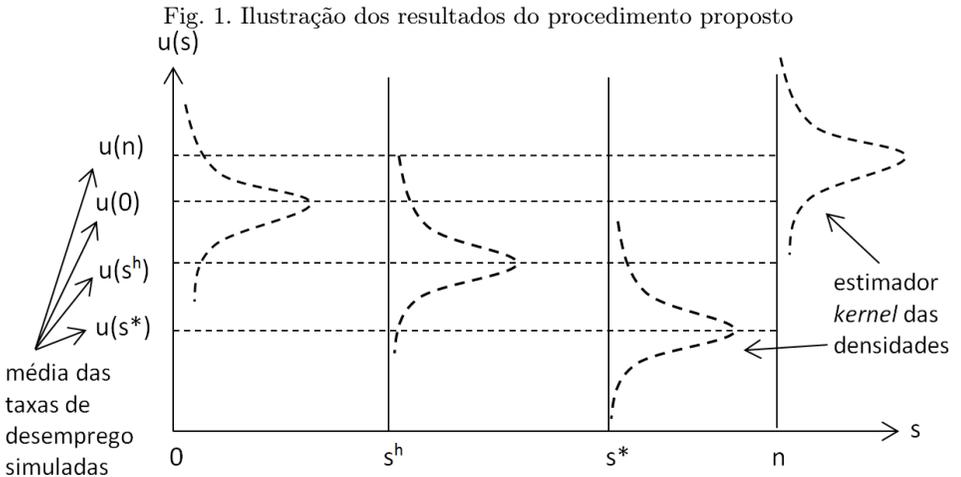
$$q_j^v = \frac{e_j^v}{n} = \frac{1}{V} \sum_{i=1}^V \frac{\exp((\hat{\beta}_0 + \sum_{k=1}^K \hat{\beta}_k d_{kj} + \hat{\xi}_j) + (\hat{\sigma}_0 z_{0i} + \sum_{k=1}^K \hat{\sigma}_k d_{kj} z_{ki}))}{1 + \sum_{r \in J} \exp((\hat{\beta}_0 + \sum_{k=1}^K \hat{\beta}_k d_{kr} + \hat{\xi}_r) + (\hat{\sigma}_0 z_{0i} + \sum_{k=1}^K \hat{\sigma}_k d_{kr} z_{ki}))} \quad (6)$$

$$q_j^s = \frac{e_j^s}{n} = \frac{1}{V} \sum_{i=1}^V \frac{\exp((\hat{\beta}_0 + \sum_{k=1}^{K-1} \hat{\beta}_k d_{kj} + \hat{\xi}_j) + (\hat{\sigma}_0 z_{0i} + \sum_{k=1}^{K-1} \hat{\sigma}_k d_{kj} z_{ki}))}{1 + \sum_{r \in J_s} \exp((\hat{\beta}_0 + \sum_{k=1}^{K-1} \hat{\beta}_k d_{kr} + \hat{\xi}_r) + (\hat{\sigma}_0 z_{0i} + \sum_{k=1}^{K-1} \hat{\sigma}_k d_{kr} z_{ki}))} \quad (7)$$

$$q_j^c = \frac{e_j^c}{n} = \frac{1}{V} \sum_{i=1}^V \frac{\exp\left(\frac{(\hat{\beta}_0 + \sum_{k=1}^{K-1} \hat{\beta}_k d_{kj} + \hat{\xi}_j) + (\hat{\sigma}_0 z_{0i} + \sum_{k=1}^{K-1} \hat{\sigma}_k d_{kj} z_{ki})}{(\hat{\beta}_K + \hat{\sigma}_K z_{Ki})}\right)}{1 + \sum_{r \in J_c} \exp\left(\frac{(\hat{\beta}_0 + \sum_{k=1}^{K-1} \hat{\beta}_k d_{kr} + \hat{\xi}_r) + (\hat{\sigma}_0 z_{0i} + \sum_{k=1}^{K-1} \hat{\sigma}_k d_{kr} z_{ki})}{(\hat{\beta}_K + \hat{\sigma}_K z_{Ki})}\right)} \quad (8)$$

Em cada cenário, considere a experimentação de t^* períodos, indexados por $t = 1, 2, \dots, t^*$. Em cada um destes períodos, seleciona-se aleatoriamente V_t postos de trabalhos vagos dentre os T postos conhecidos. E em cada cenário e período experimentado coletam-se as taxas de desemprego estimadas por $u_t = 1 - \sum_j q_{jt}$. O processo gerará três vetores de taxas de desemprego simuladas (cada um de dimensão t^*). E sobre cada um destes vetores pode-se construir um estimador *kernel* da função de densidade das taxas de desemprego dado o uso da rede social.

A Figura 1 ilustra no plano cartesiano a taxa de desemprego em função do tamanho da rede social, $u(s)$, onde $0 \leq s \leq n$ (n é o número total de trabalhadores). O tamanho ótimo da rede, s^* , é desconhecido. Mas a simulação do cenário de referência reflete o tamanho atual da rede, s^h , e os outros cenários refletem os casos extremos ($s = 0$ e $s = n$). A simulação de t^* processos de *matching* em cada cenário permitirá a construção e visualização das densidades das taxas de desemprego resultantes em cada cenário.



Elaboração própria.

O modelo CAZ prevê que a taxa de desemprego, em estado estacionário, é tal que: $u(0) < u(s^h) < u(n)$ ou $u(n) < u(s^h) < u(0)$. Então, se em uma inspeção

visual acompanhada de testes de diferenças de médias as densidades apontarem que $\mathbb{E}(u(0)) < \mathbb{E}(u(s^h)) < \mathbb{E}(u(n))$ ou $\mathbb{E}(u(n)) < \mathbb{E}(u(s^h)) < \mathbb{E}(u(0))$, as simulações corroborarão com o previsto no modelo. Ou seja, isto indicaria que existe um tamanho ótimo de rede para se minimizar a taxa de desemprego da economia. Qualquer outro ordenamento nestas desigualdades apontaria contra a validade do modelo.

4. Base de Dados

Para o exercício empírico foram usadas as informações do *The European Community Household Panel* (ECHP) para Portugal, nos anos 1998 a 2001. Em cada ano foram entrevistados cerca de seis mil trabalhadores (empregados e desempregados).

Considerando que o efeito das redes sociais pode ser diferente para empregos que exigem diferentes níveis de escolaridade (detalhes em Ioannides e Loury 2004), estratificou-se a amostra para trabalhadores de níveis fundamental, médio e superior. As estatísticas descritivas são apresentadas no quadro da Figura 2, adiante.

Note que a taxa de desemprego observada na amostra é maior para trabalhadores de escolaridade fundamental e média (cerca de 5% a 7%, dependendo do ano) do que para trabalhadores com nível superior (cerca de 2% a 3%). Os números apresentados estão em consonância com outros registros do Instituto Nacional de Estatísticas de Portugal, independentes do ECHP, e analisados em Portugal (2003).

No quadro pode-se observar que, para o nível de escolaridade fundamental, perto de 60% dos trabalhadores são homens. Ao passo que este número se reduz para perto de 50% e 40% para os níveis médio e superior, respectivamente. E, independentemente do nível de escolaridade, a proporção de trabalhadores casados é maior que a dos outros estados civis.

As linhas seguintes do quadro apresentam a proporção de trabalhadores observados por faixa etária. Para o nível fundamental de escolaridade a distribuição de idades é relativamente homogênea entre 20 e 50 anos. Ao passo que para os outros níveis se observam mais trabalhadores na faixa entre 26 e 30 anos de idade.

Em relação ao uso da rede social, o questionário do ECHP possui a seguinte pergunta para quem se declara empregado: *para encontrar trabalho você...*

- a) aplicou um currículo diretamente ao empregador;
- b) encontrou a vaga em anúncio de jornal ou assemelhado;
- c) encontrou a vaga através de uma agência de empregos;
- d) recebeu a indicação de amigos ou familiares; ou,
- e) começou seu próprio negócio ou se juntou aos negócios da família.

A última linha do quadro apresenta a proporção de respostas na opção d) (entendendo que a pessoa usou sua rede social para encontrar seu emprego).

Quanto ao uso da rede, note que de 29% a 34% dos trabalhadores com nível de

Fig. 2. Estatísticas descritivas da amostra por ano e escolaridade do trabalhador observado. Fonte: ECHP/Portugal

Variável / Ano	Ensino Fundamental				Ensino Médio				Ensino Superior			
	1998	1999	2000	2001	1998	1999	2000	2001	1998	1999	2000	2001
Taxa de desemprego (%)	7,8	6,4	5,4	5,2	7,0	6,7	5,6	6,1	2,5	3,0	2,3	3,1
% de homens	59,7	59,8	59,4	59,6	50,7	51,0	49,8	49,8	41,8	39,3	39,3	38,3
% de casados	69,7	70,0	70,2	71,2	56,6	55,5	56,0	57,9	63,0	63,4	62,9	61,0
% de solteiros	23,2	22,6	22,2	22,0	37,7	38,4	40,2	38,7	30,6	30,0	30,9	32,7
% de separados/divorciados	3,8	3,9	4,0	4,2	4,9	5,6	3,6	3,2	5,5	5,7	5,0	5,1
% viúvos	3,3	3,5	3,6	2,6	0,8	0,5	0,2	0,2	0,8	1,0	1,1	1,1
% menos de 20 anos	5,3	4,8	4,9	4,8	2,0	2,3	2,4	2,0	0,5	0,1	0,1	0,1
% de 20 a 25 anos	11,5	11,4	10,3	9,5	22,6	21,5	20,1	20,8	8,2	6,2	6,5	6,7
% de 26 a 30 anos	11,3	11,5	12,5	12,8	23,6	26,5	27,6	25,7	22,9	26,6	27,2	27,5
% de 31 a 35 anos	9,8	9,5	10,0	10,0	15,2	13,4	13,8	15,3	17,1	16,6	17,7	17,2
% de 36 a 40 anos	11,2	11,3	10,9	10,6	13,4	13,7	14,3	14,0	13,9	13,8	14,0	13,0
% de 41 a 45 anos	11,6	11,3	11,2	11,3	9,7	8,4	8,2	9,3	11,9	11,0	10,1	9,3
% de 46 a 50 anos	10,5	10,8	11,0	11,6	6,3	6,3	6,4	5,7	10,0	10,7	8,6	8,9
% de 51 a 55 anos	9,2	9,2	9,0	9,7	4,1	4,2	4,6	4,0	8,0	8,2	8,3	9,6
% de 56 a 60 anos	8,0	7,9	8,0	7,6	1,7	1,9	1,2	2,0	4,8	4,7	5,5	4,3
% de 61 a 65 anos	6,7	6,9	6,5	6,8	0,9	1,1	1,1	1,0	1,3	1,4	1,2	2,1
% de 65 a 70 anos	3,5	3,6	3,9	3,6	0,4	0,6	0,4	0,1	0,7	0,6	0,7	1,0
% mais de 70 anos	1,4	1,8	1,9	1,8	0,1	0,1	0,1	0,1	0,6	0,1	0,1	0,4
% usou a rede social	29,4	31,4	32,8	33,9	30,2	33,8	34,8	34,1	17,2	18,1	18,9	18,0

Elaboração própria.

escolaridade fundamental ou médio encontraram emprego pela indicação de amigos ou familiares. E este número é menor para o nível superior de escolaridade – cerca de 17% a 18%.

Nesse sentido, Datcher (1983) e Wahba e Zenou (2005) já haviam documentado que o uso da rede é menor para maiores níveis de escolaridade. A explicação para isto seria que, em empregos que exigem maior escolaridade, o preparo técnico e a experiência são mais importantes que a indicação.

Sobre o tratamento da amostra para a implementação do *mixed-logit*, ao invés de se trabalhar com um painel (de quatro anos e até 2^K unidades *cross-section*) optou-se em agregar todos os anos como um único período por três razões. A primeira é uma consequência do uso do sistema de numeração posicional. Para perceber isso, imagine $K = 10$. Então o pesquisador poderia encontrar até 1024 tipos de trabalhadores. Mas já nas primeiras tentativas de se operacionalizar o exercício percebeu-se que muitos tipos deixam de aparecer em anos subsequentes. O que acaba por gerar um grande número de *missings* para se montar um painel. Mas esse número é radicalmente reduzido quando se considera todos os anos em um único período.

É claro que o número de *missings* não gera uma impossibilidade, uma vez que existem estimadores específicos para esses casos que podem ser processados dentro da rotina computacional do *mixed-logit*. Todavia, estes estimadores necessitam que as variáveis explicativas mudem de valor entres os períodos analisados. Então a segunda razão de se optar em usar todos os anos como um único período é de cunho econométrico. Pois para que o sistema proposto na seção anterior seja aplicável,

todas as variáveis explicativas inseridas na computação do modelo são *dummies* que caracterizam os tipos, logo não podem mudar de valor.

A terceira razão é de origem numérica. Pois a rotina usada para se estimar os parâmetros do *mixed-logit* envolve uma *contração*⁵ que computaria, nesse exemplo, um vetor de $1024 \times 4 = 4096$ entradas (muitas delas *missing*). O que acaba por envolver números extremamente pequenos, demandando critérios para aceitação de convergência da contração ainda menores e aumentando potenciais erros de simulação.

Sobre a determinação das K características a se usar, existe um dilema: se K for pequeno demais, o exercício fica irrealista; e se K for grande demais, o exercício fica numericamente inviável. Então esta escolha deve ser feita com parcimônia.

No exercício se particionará os tipos através do gênero (uma *dummy* indicando sexo masculino), do estado civil (*dummies* indicando os estados casado, separado/divorciado e viúvo), das faixas etárias em 11 classes (*dummies* indicando espaçamentos de cinco anos na idade, começando em 20 e terminando com uma faixa para idades superiores a 70 anos) e se declarou ter encontrado emprego pela rede social (uma *dummy* indicando que recebeu indicação da rede). Assim, usa-se $K = 16$. Onde o *tipo base*, na terminologia usada na seção anterior, é uma mulher, solteira, com menos de 20 anos de idade que não usou sua rede social para encontrar o emprego.

Após a computação das características dos trabalhadores, a base de dados permitiu a observação de 171 diferentes tipos com nível fundamental, 104 com nível médio e 97 com nível superior. Estes valores são o J do modelo proposto. E é interessante notar que, neste caso, o número potencial de diferentes tipos computados com o sistema de numeração posicional era de $2^{16} = 65536$. Então, na prática, se observa um número substancialmente menor que o potencial.

5. Resultados Estimados

A equação (5) foi estimada com o algoritmo proposto por Berry e Pakes (2007). Usou-se $T = 1000$, e a função objetivo minimizada foi a soma dos quadrados de ξ_j (o índice que agrega as características do trabalhador que não puderam ser captadas pelo sistema de numeração posicional).⁶

Como discutido em Train (2009), os parâmetros do *mixed logit* não possuem uma interpretação simples. Pois estão relacionados com as idiosincrasias da função de escolha da firma, dada pela equação (2): $E_{j\tau} = \Omega_{j\tau} + \varepsilon_{j\tau}$; onde $\Omega_{j\tau} = \beta_{0\tau} + \sum_{k=1}^K \beta_{k\tau} d_{kj} + \xi_j$.

Por construção, normaliza-se o termo $\Omega_{u\tau} = 0$ se a firma opta por não preencher a vaga.⁷ E quanto maior o valor $\beta_{k\tau} = \bar{\beta}_k + z_{k\tau} \sigma_k$, maior será a chance do trabalhador

⁵ Ver a discussão da p. 322 de Train (2009).

⁶ O vetor z foi formado com a semente aleatória 987654321.

⁷ Mesmo com a normalização, a escala das escolhas não é centrada em zero. Pois $\mathbb{E}(\varepsilon_{j\tau}) = -\gamma$, onde γ é a constante de Euler-Mascheroni (0,57721...). E a probabilidade da firma manter uma vaga desocupada

com a característica k ser escolhido. Então, quanto maior o peso médio ($\bar{\beta}_k$), e menor o desvio (σ_k), é mais provável que um trabalhador consiga se empregar. Mas, devido as idiossincrasias do trabalhador (ξ_j) e da firma ($z_{k\tau}$ e $\varepsilon_{j\tau}$), menores pesos não implicarão necessariamente que um trabalhador não incorrerá em *matching*.

Considerando apenas os trabalhadores com nível de escolaridade fundamental, a Figura 3 apresenta o quadro com os valores dos parâmetros estimados para a equação (5). Também são apresentados os respectivos desvios padrão estimados para estes parâmetros e as estatísticas t .

Fig. 3. Estimativas da equação (5) para trabalhadores com nível fundamental de escolaridade. Parâmetros da média ($\bar{\beta}$) e do desvio (σ) dos pesos da função escolha das firmas. $J = 171$

Covariada	Parâmetro da média			Parâmetro do desvio		
	Estimativa	Desvio padrão	Estatística t	Estimativa	Desvio padrão	Estatística t
<i>Constante</i>	-2,9782	0,3991	-7,4627	1,7276	0,3866	4,4686
<i>Dummy homem</i>	-0,0383	0,1724	-0,2223	0,7798	0,0445	17,5213
<i>Dummy casado</i>	1,3304	0,2349	5,6638	0,3840	0,0301	12,7495
<i>Dummy separado/divorciado</i>	-1,5928	0,2405	-6,6224	0,5135	0,1419	3,6200
<i>Dummy viúvo</i>	-2,2476	0,2598	-8,6501	0,9847	0,1315	7,4887
<i>Dummy 20 a 25 anos</i>	0,6421	0,4816	1,3330	1,9903	0,2220	8,9664
<i>Dummy 26 a 30 anos</i>	0,2956	0,4617	0,6403	1,9857	0,3253	6,1034
<i>Dummy 31 a 35 anos</i>	0,8161	0,4568	1,7865	1,4595	0,1123	12,9919
<i>Dummy 36 a 40 anos</i>	0,4062	0,4627	0,8799	2,1235	0,2871	7,3951
<i>Dummy 41 a 45 anos</i>	0,7824	0,4569	1,7126	1,5912	0,1524	10,4429
<i>Dummy 46 a 50 anos</i>	0,3212	0,4567	0,7031	2,0057	0,4151	4,8316
<i>Dummy 51 a 55 anos</i>	0,3778	0,4617	0,8183	1,7296	0,2354	7,3463
<i>Dummy 56 a 60 anos</i>	-0,2179	0,4628	-0,4709	2,1027	0,9691	2,1697
<i>Dummy 61 a 65 anos</i>	-0,1053	0,4636	-0,2270	1,5056	0,0943	15,9725
<i>Dummy 65 a 70 anos</i>	-1,5369	0,4711	-3,2621	2,2264	0,8430	2,6412
<i>Dummy mais de 70 anos</i>	-1,0088	0,4834	-2,0870	1,5903	0,2334	6,8147
<i>Dummy rede social</i>	-1,2963	0,1729	-7,4963	1,4750	0,2307	6,3945

Elaboração própria.

Uma interpretação dos parâmetros para a característica de gênero é a seguinte: em 95% dos casos, as firmas dão um peso ($\beta_{k\tau}$) que se encontra no intervalo $[-1, 56; 1, 49]$ para o fato do trabalhador ser homem.⁸ Logo, existem vagas onde se pondera positivamente esta característica, e vagas onde o fato de ser homem pesa negativamente.

Quanto ao estado civil, em 95% dos casos o peso que uma firma dá as condições de casado, separado/divorciado e viúvo encontram-se, respectivamente, nos intervalos $[0, 58; 2, 08]$, $[-2, 06; -0, 59]$ e $[-4, 18; -0, 38]$. E em relação as faixas etárias, observa-se uma ponderação média positiva para os trabalhadores de até 55 anos de idade. E uma ponderação média negativa para idades superiores aos 56 anos.

é: $q_{u\tau} = 1/(1 + \sum_{r \in J} \exp(\Omega_{r\tau}))$.

⁸ O intervalo explora as características da distribuição normal, e equivale a $\bar{\beta} \pm 1,96 \times \sigma = -0,0383 \pm 1,96 \times 0,7798$.

E quanto ao uso da rede social, em 95% dos casos o peso que cada firma dá ao trabalhador indicado por alguém encontra-se no intervalo $[-4, 19; 1, 59]$. Uma interpretação deste intervalo é que um mesmo trabalhador, ao ser indicado pela rede, pode receber uma ponderação positiva para algumas vagas e negativa para outras.

Fig. 4. Estimativas da equação (5) para trabalhadores com nível médio de escolaridade. Parâmetros da média ($\bar{\beta}$) e do desvio (σ) dos pesos da função escolha das firmas. $J = 104$

Covariada	Parâmetro da média			Parâmetro do desvio		
	Estimativa	Desvio padrão	Estatística <i>t</i>	Estimativa	Desvio padrão	Estatística <i>t</i>
<i>Constante</i>	-3,0103	0,3823	-7,8739	1,4967	3,2358	0,4626
<i>Dummy homem</i>	-0,0313	0,1940	-0,1612	0,2230	0,0666	3,3459
<i>Dummy casado</i>	0,7567	0,2291	3,3034	0,0204	0,0741	0,2754
<i>Dummy separado/divorciado</i>	-1,6047	0,2591	-6,1940	0,2733	0,3690	0,7407
<i>Dummy viúvo</i>	-1,8570	0,4869	-3,8136	1,2247	1,8102	0,6765
<i>Dummy 20 a 25 anos</i>	1,5143	0,4750	3,1882	2,0697	2,8819	0,7182
<i>Dummy 26 a 30 anos</i>	1,7547	0,4591	3,8217	2,0681	1,2020	1,7206
<i>Dummy 31 a 35 anos</i>	1,4219	0,4539	3,1327	1,2176	0,5373	2,2659
<i>Dummy 36 a 40 anos</i>	0,6146	0,4749	1,2942	2,3187	1,9715	1,1761
<i>Dummy 41 a 45 anos</i>	0,5422	0,3579	1,5149	1,5571	0,9045	1,7215
<i>Dummy 46 a 50 anos</i>	-0,2144	0,4670	-0,4592	2,0188	1,7208	1,1732
<i>Dummy 51 a 55 anos</i>	-0,3812	0,4595	-0,8296	1,5966	0,3720	4,2925
<i>Dummy 56 a 60 anos</i>	-1,5623	0,5177	-3,0181	2,1856	2,1503	1,0164
<i>Dummy 61 a 65 anos</i>	-0,9968	0,5345	-1,8648	1,4040	0,6411	2,1900
<i>Dummy 65 a 70 anos</i>	-3,1619	0,6522	-4,8477	2,3479	2,9055	0,8081
<i>Dummy mais de 70 anos</i>	-1,8962	0,7704	-2,4614	1,3259	3,3863	0,3916
<i>Dummy rede social</i>	-0,8655	0,1901	-4,5533	1,3125	0,1596	8,2247

Elaboração própria.

A figura 4 apresenta o quadro com os valores estimados para os trabalhadores com nível médio de escolaridade. Para as variáveis de gênero e estado civil, os parâmetros apontam um comportamento semelhante aos apresentados para o nível fundamental. Não há uma ponderação média significativa diferenciando o gênero,⁹ e há uma ponderação positiva para os trabalhadores casados.

O comportamento semelhante também ocorre para as características de idade, com ponderações médias negativas para as faixas etárias mais elevadas. E quanto ao uso da rede social, em 95% dos casos o peso que cada firma dá ao trabalhador indicado por alguém encontra-se no intervalo $[-3, 41; 1, 95]$.

A figura 5 apresenta o quadro com os valores dos parâmetros estimados para os trabalhadores com nível superior. E o mesmo padrão observado nos outros níveis de escolaridade se repete aqui: não há diferença significativa para o peso médio em relação ao gênero; há uma ponderação média maior para trabalhadores casados; e há uma ponderação média menor para trabalhadores com idade acima de 56 anos. E quanto ao uso da rede social, em 95% dos casos o peso que cada firma dá ao trabalhador indicado por alguém encontra-se no intervalo $[-3, 01; 1, 46]$.

⁹ A 5% de significância.

Fig. 5. Estimativas da equação (5) para trabalhadores com nível superior de escolaridade. Parâmetros da média ($\bar{\beta}$) e do desvio (σ) dos pesos da função escolha das firmas. $J = 97$

Covariada	Parâmetro da média			Parâmetro do desvio		
	Estimativa	Desvio padrão	Estatística t	Estimativa	Desvio padrão	Estatística t
<i>Constante</i>	-3,5260	0,6915	-5,0990	0,0166	5,9007	0,0028
<i>Dummy homem</i>	-0,2793	0,2017	-1,3846	0,2846	1,1809	0,2410
<i>Dummy casado</i>	0,8473	0,2329	3,6378	0,0153	0,0615	0,2489
<i>Dummy separado/divorciado</i>	-0,8698	0,3043	-2,8586	0,2796	0,9491	0,2946
<i>Dummy viúvo</i>	-4,0203	0,4114	-9,7727	2,4429	1,4016	1,7429
<i>Dummy 20 a 25 anos</i>	0,9723	0,7629	1,2746	1,8596	1,7014	1,0930
<i>Dummy 26 a 30 anos</i>	2,1294	0,7503	2,8382	2,6805	1,7177	1,5605
<i>Dummy 31 a 35 anos</i>	2,9685	0,7666	3,8725	0,3202	1,8893	0,1695
<i>Dummy 36 a 40 anos</i>	0,3369	0,7005	0,4809	3,6978	7,3726	0,5016
<i>Dummy 41 a 45 anos</i>	0,4482	0,7428	0,6034	3,5632	1,6755	2,1266
<i>Dummy 46 a 50 anos</i>	0,1897	0,7547	0,2514	2,7646	1,1324	2,4413
<i>Dummy 51 a 55 anos</i>	0,7124	0,7479	0,9525	0,8797	0,4028	2,1842
<i>Dummy 56 a 60 anos</i>	-2,3342	0,7611	-3,0670	3,8948	1,5567	2,5020
<i>Dummy 61 a 65 anos</i>	-1,0957	0,7891	-1,3885	0,1133	3,6863	0,0307
<i>Dummy 65 a 70 anos</i>	-2,3388	0,8424	-2,7763	2,6295	2,0005	1,3144
<i>Dummy mais de 70 anos</i>	-1,0993	0,8626	-1,2744	0,0518	0,0310	1,6689
<i>Dummy rede social</i>	-1,0821	0,2247	-4,8160	1,0199	0,0581	17,5436

Elaboração própria.

Tendo em mãos os parâmetros estimados para o *mixed-logit*, o próximo passo é transformar parte dos postos de trabalho em vagas e simular como os processos de *matching* são afetados pela rede social. Operacionalmente, faz-se a experimentação de $t^* = 100$ períodos, e em cada período seleciona-se aleatoriamente $V_t = 200$ postos diferentes dentre os $T = 1000$ usados na estimação de (5).

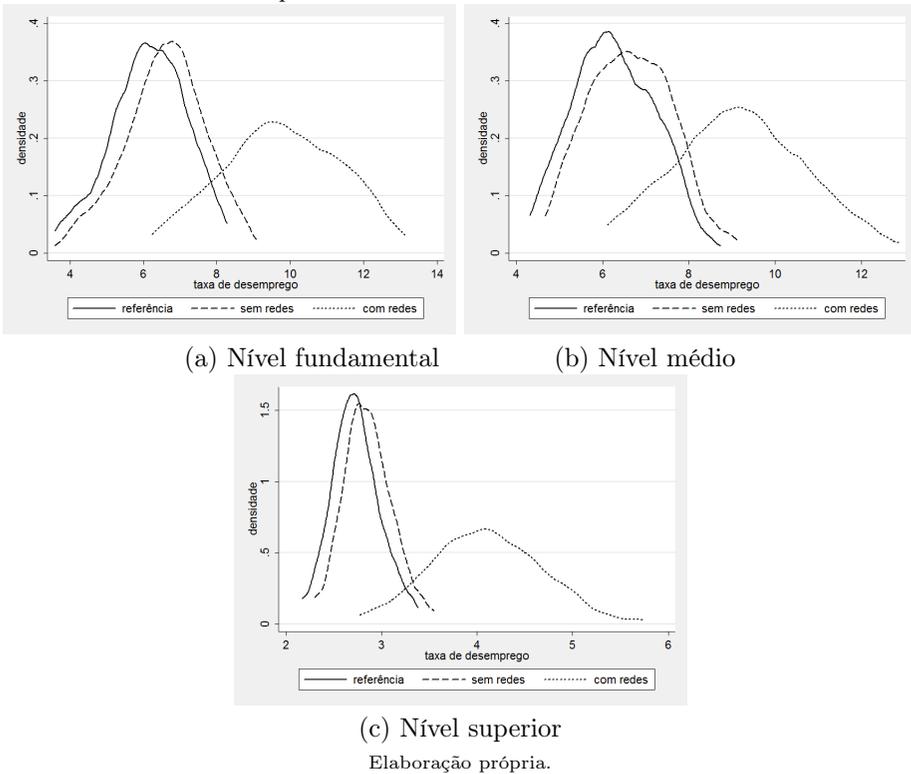
São considerados três cenários de preenchimento das vagas, e em cada um são computadas as taxas de desemprego por $u_t = 1 - \sum_j q_{jt}$, onde q_{jt} é resultante das equações (6), (7) ou (8), dependendo do caso. Há um cenário referencial, quando o uso das redes não é alterado – apenas simulam-se novos processos de *matching* para as V_t vagas. E os outros cenários são aqueles onde nenhum e todo trabalhador, respectivamente, faz uso da rede.

O exercício é feito para cada estratificação de escolaridade. E sobre cada um dos vetores contendo as taxas de desemprego simuladas por período, cenário e escolaridade, são construídos estimadores *kernel* da função de densidade. Este procedimento permite uma inspeção visual dos resultados das simulações.

Observe, primeiramente, as densidades das simulações para o nível fundamental – item (a) da Figura 6. Da esquerda para a direita, as curvas ilustram os cenários de referência e onde nenhum e todo trabalhador faz uso da rede social. No cenário de referência observam-se taxas de desemprego com densidade semelhante ao cenário onde não se usa a rede social (embora centradas à esquerda, indicando um cenário mais favorável aos trabalhadores).

Por outro lado, essa densidade se modifica quando todos usam a rede. Sua variância fica maior e a frequência é maior em taxas mais altas. Ou seja, é um cenário com taxas de desemprego maiores e mais irregulares.

Fig. 6. Gráficos do estimador *kernel* das funções densidade das taxas de desemprego (%) simuladas em cada cenário para cada nível de escolaridade



O itens (b) e (c) da Figura 6 apresentam as densidades da taxa de desemprego para os níveis médio e fundamental, respectivamente, para cada cenário. Pela inspeção visual, note-se um padrão para todos os níveis de escolaridade, pontuados da seguinte forma:

- 1) as taxas de desemprego menores são mais frequentes no cenário de referência – onde alguns trabalhadores usam a rede social e outros não;
- 2) a variância das taxas de desemprego do cenário de referência é semelhante a do cenário onde ninguém usa a rede; e,
- 3) as taxas de desemprego do cenário onde todos usam a rede apresenta maior variância.

Em relação a isso, algumas conjecturas podem ser feitas. O primeiro padrão corrobora com o previsto no modelo teórico, que aponta a existência de uma tamanho ótimo de rede social para se minimizar a taxa de desemprego. Como há dois cenários extremos, nenhum deles deve refletir o ponto ótimo. E como o cenário de referência é intermediário aos extremos, suas taxas de desemprego devem ser mais próximas do que seria observado no cenário ótimo. Escrito de outra forma, o primeiro padrão reflete a seguinte ordem de expectativas: $\mathbb{E}(u^*) \leq \mathbb{E}(u^v) \leq$

$\mathbb{E}(u^s) \leq \mathbb{E}(u^c)$; para as taxas de desemprego dos cenários ótimo, de referência e quanto ninguém usa e todos usam a rede social, respectivamente.¹⁰

Quanto ao segundo padrão, é possível que o número de conexões sociais dos trabalhadores observados seja pequeno em relação ao ótimo. Ou seja, na terminologia do modelo teórico, as densidades parecem refletir que $0 < s^v < s^* < n$, onde s^v é o tamanho (representativo) das redes sociais dos trabalhadores observados. E s^v está mais próximo de zero do que de s^* , por isso observa-se a semelhança entre as densidades.

O terceiro padrão é a repercussão do aumento da variância em cada probabilidade do trabalhador do tipo j encontrar emprego (q_j). E este aumento, por sua vez, decorre do fato que todos os trabalhadores recebem indicação da rede. Logo o peso β_{Ki} , que é aleatório, aparece para todos na função de escolha da firma (enquanto nos outros cenários aparecia apenas para alguns ou nunca aparecia). Então aumenta-se a variância de todo q_j , aumentando-se a variância de $u_t = 1 - \sum_j q_{jt}$.

A Figura 7, apresenta as estatísticas que descrevem as densidades estimadas. Os valores são usados em testes de diferenças de média entre cada cenário e escolaridade.

Fig. 7. Média, desvio padrão, máximo e mínimo observados nas taxas de desemprego (%) simuladas por cenário e nível de escolaridade

Estatística / Cenário	Ensino Fundamental			Ensino Médio			Ensino Superior		
	referência	sem redes	com redes	referência	sem redes	com redes	referência	sem redes	com redes
Média	6,15	6,58	9,76	6,25	6,63	9,20	2,73	2,86	4,10
Desvio Padrão	1,02	1,08	1,58	0,95	0,96	1,51	0,25	0,26	0,58
Máximo	8,28	9,08	13,14	8,74	9,13	12,86	3,38	3,55	5,74
Mínimo	3,59	3,59	6,24	4,32	4,68	6,12	2,17	2,30	2,77

Elaboração própria.

Comparando os cenários de referência em relação aos ausentes de rede social, as diferenças de média das taxas de desemprego simuladas são $-0,430$, $-0,387$ e $-0,131$ para os casos das escolaridades fundamental, média e superior, respectivamente. Os intervalos de confiança (95%) para estas diferenças, considerando variâncias diferenciadas, são de $[-0,723; -0,137]$, $[-0,653; -0,121]$ e $[-0,201; -0,059]$. Como o zero não está contido em nenhum dos intervalos, conclui-se que em média as taxas de desemprego do cenário de referência são menores que no cenário onde nenhum trabalhador usa a rede.

Comparando os cenários de referência em relação aos que todos usam a rede social, as diferenças de média são $-3,618$, $-2,956$ e $-1,372$ para os casos das escolaridades fundamental, média e superior, respectivamente. Os intervalos de confiança (95%), considerando variâncias diferenciadas, são de $[-3,240; -3,995]$, $[-2,602; -3,309]$ e $[-1,245; -1,498]$. Assim, também conclui-se que em média as taxas de desemprego do cenário de referência são menores que no cenário onde todos os trabalhadores usam a rede.

¹⁰ Como na ilustração da Figura 1.

Os resultados dos testes corroboram que $\mathbb{E}(u^v) < \mathbb{E}(u^s) < \mathbb{E}(u^c)$. Pois com mais frequência as taxas de desemprego são maiores nos cenários extremos. Logo, não se pode rejeitar a tese de que existe um tamanho ótimo de redes sociais para se minimizar desemprego.

6. Considerações Finais

Segundo Calvó-Armengol (2004), Calvó-Armengol e Zenou (2005) e Wahba e Zenou (2005), existe um tamanho ótimo de rede social capaz de minimizar a taxa de desemprego da economia. Sendo assim, as formas das redes poderiam explicar parte de uma inflexibilidade em se reduzir desemprego. Mas a tese é difícil de se testar empiricamente, uma vez que é difícil de se observar o tamanho das redes sociais.

O artigo propôs um procedimento para simular as escolhas das firmas de como preencher um posto de trabalho com determinados tipos de trabalhadores. Nas simulações os desempregados podem (ou não) fazer uso da sua rede social para melhorar a empregabilidade. Aplicou-se a proposta em uma base dados para Portugal, e os resultados estimados corroboraram com a tese do tamanho ótimo de rede social.

O procedimento proposto pode ser aplicado em outras bases de dados de pesquisas de emprego e desemprego. Então, se os resultados encontrados aqui forem observados em outros exercícios haverá maiores evidências de que existe tal tamanho ótimo de rede social.

Uma possível inovação na proposta, a se realizar em trabalhos futuros, é simular o tamanho da rede social de forma contínua. Isto é, mais do que usar os cenários extremos e o cenário de referência, como foi feito aqui, talvez seja possível simular vários tamanhos de rede explorando o fato de que diferentes tipos de trabalhadores usam a rede com intensidades diferentes.

Por fim, observa-se que o procedimento proposto tem um potencial uso na avaliação de políticas públicas. Pois as simulações tratam do uso (ou não) das redes sociais e seu impacto na taxa de desemprego. E uma rápida alteração na técnica pode, por exemplo, simular a aplicação (ou não) de determinado programa de treinamento e seu impacto na taxa de desemprego. Este também é um possível trabalho futuro.

Referências bibliográficas

- Bartlett, R. & Miller, T. (1985). Executive compensation: Female executives and networking. *The American Economic Review*, 75:266–270.
- Berry, S. & Pakes, A. (2007). The pure characteristics demand model. *International Economic Review*, 48:1193–1225.
- Blanchard, O., Diamond, P., Hall, R., & Yellen, J. (1989). The Beveridge curve. *Brookings Papers on Economic Activity*, 1:1–76.

- Blanchard, O. & Portugal, P. (2001). What hides behind an unemployment rate: Comparing Portuguese and U.S. labor markets. *The American Economic Review*, 91:187–207.
- Calvó-Armengol, A. (2004). Job contact networks. *Journal of Economic Theory*, 115:191–206.
- Calvó-Armengol, A. & Zenou, Y. (2005). Job matching, social network and word-of-mouth communication. *Journal of Urban Economics*, 57:500–522.
- Corcoran, M., Datcher, L., & Duncan, G. (1980). Most workers find jobs through word of mouth. *Monthly Labor Review*, 103:33–35.
- Datcher, L. (1983). The impact of informal networks on quit behavior. *The Review of Economics and Statistics*, 65:491–495.
- Holzer, H. (1988). Search method used by unemployed youth. *Journal of Labor Economics*, 6:1–20.
- Ioannides, Y. & Loury, L. (2004). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, 42:1056–1093.
- McFadden, D. (2001). Economic choices. *The American Economic Review*, 91:351–378.
- McFadden, D. & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15:447–470.
- Montgomery, J. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *The American Economic Review*, 81:1408–1418.
- Montgomery, J. (1994). Weak ties, employment, and inequality: An equilibrium analysis. *The American Journal of Sociology*, 99:1212–1236.
- Petrongolo, B. & Pissarides, C. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic Literature*, 39:390–431.
- Pissarides, C. (2000). *Equilibrium Unemployment Theory*. The MIT Press, second edition.
- Portugal, P. (2003). Emprego, salários e condições de trabalho. In *Portugal Social*, pages 91–112. Instituto Nacional de Estatística.
- Rees, A. (1966). Information networks in labor markets. *The American Economic Review*, 56:559–566.
- Scheinerman, E. (2005). *Mathematics: A Discrete Introduction*. Brooks Cole.
- Topa, G. (2001). Social interactions, local spillovers and unemployment. *Review of Economic Studies*, 68:261–295.
- Train, K. (2009). *Discrete Choice Methods With Simulation*. Cambridge University Press.
- Wahba, J. & Zenou, Y. (2005). Density, social networks and job search methods: Theory and application to Egypt. *Journal of Development Economics*, 78:443–473.