

## 2 Base de Dados

Os dados utilizados neste trabalho foram obtidos na Pesquisa Nacional por Amostra de Domicílios (PNAD), realizada anualmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Esta pesquisa cobre todo o país, com exceção da área rural dos estados da antiga Região Norte (Rondônia, Acre, Amazonas, Roraima, Pará e Amapá). São investigadas, basicamente, informações relativas a características gerais da população, educação, trabalho, rendimento e condições de habitação. Nesta seção, são apresentados os detalhes da construção da base de dados e as suas características.<sup>7</sup>

### 2.1. Seleção da Amostra e Definições Básicas

Foram utilizadas as pesquisas que ocorreram entre os anos de 1981 e 2001.<sup>8</sup> Além disso, foi feita uma seleção de domicílios onde tanto o chefe quanto o cônjuge tinham a idade limitada entre 20 e 29 anos, inclusive. Esta restrição se deve ao fato da riqueza para os jovens ser pouco determinada pelas suas escolhas passadas. A riqueza exógena permite contornar um importante problema de endogeneidade. O que se pretende mostrar é que a riqueza determina a escolha ocupacional, mas esta, por sua vez, define a renda e através da renda impacta a riqueza. No entanto, se a riqueza é exógena, pode-se ignorar esse segundo efeito.

Embora a pesquisa não reporte diretamente a riqueza dos indivíduos, foi possível obter uma *proxy* para a riqueza dos domicílios, discutida na próxima seção. Alguns domicílios continham mais de uma família, não sendo possível saber a contribuição de cada uma delas para a riqueza do domicílio. Alternativamente a repartir a riqueza pelo número de famílias ou atribuir a riqueza a uma família específica, a solução natural foi excluir da amostra os domicílios

---

<sup>7</sup> Todas as figuras e tabelas deste trabalho foram feitas usando essa base de dados.

<sup>8</sup> A pesquisa foi interrompida para a realização do Censo Demográfico em 1991 e 2000 e não ocorreu em 1994 por razões excepcionais.

com essa característica. Somente quando o domicílio é composto por uma família principal e outras famílias que tenham como chefe um empregado doméstico do domicílio, pode-se identificar que família é de fato a detentora da riqueza domiciliar. Nesses casos, a riqueza foi atribuída a ela, tendo sido excluídas as demais.

O levantamento das características do domicílio só é efetuado na PNAD quando o domicílio é particular permanente. Domicílios particulares são aqueles domicílios destinados para habitação de um grupo de pessoas cujo relacionamento é ditado por parentesco, dependência doméstica ou normas de convivência. Ou seja, o relacionamento não se restringe ao cumprimento de normas administrativas. Os domicílios particulares são classificados como permanentes quando se localizam em casas, apartamentos ou cômodos e são destinados à moradia. Mais uma vez, como se depende das características dos domicílios para a construção da *proxy* para riqueza, a amostra utilizada neste trabalho se restringe a famílias que residem neste tipo de domicílio.

Faz-se necessário, ainda, esclarecer o que caracteriza um chefe de família na pesquisa, como está definido o conceito de trabalho e que tipo de pessoa será considerada como empresário. O chefe da família é a pessoa responsável pela família ou que assim é considerada pelos demais membros. Nesta pesquisa são considerados ocupados os indivíduos que exercem alguma atividade remunerada ou alguma atividade não remunerada por pelo menos quinze horas semanais. Este trabalho considera como empresário as pessoas cuja ocupação, segundo a PNAD, é ser um empregador.<sup>9</sup>

O cálculo dos anos de estudo é feito em função do último grau e série concluídos. O primeiro ano de estudo equivale à conclusão da primeira série do primeiro grau. Cada série concluída corresponde ao acréscimo de mais um ano de estudo. O primeiro segmento do ensino fundamental (1º a 4º série do primeiro grau) leva quatro anos para ser concluído. De 5 a 8 anos de estudo conclui-se o

---

<sup>9</sup> Não serão considerados empresários, portanto, os trabalhadores por conta-própria. No Brasil, a atividade por conta-própria tem sido associada à informalidade e pobreza. Segundo Ferreira, Lanjouw e Neri (2000), a maioria dos chefes de família que estão abaixo da linha de pobreza é trabalhador por conta-própria. Ou seja, o trabalho por conta-própria é a posição na ocupação que mais contribui para a pobreza no Brasil. Para uma maior compreensão deste tipo de ocupação ver Neri (1999a e 1999b). Estes dois trabalhos se referem aos trabalhadores por conta-própria como empresários, mas apontam a existência de diferenças entre os empresários por conta-própria e aqueles empresários que geram empregos.

segundo segmento do ensino fundamental. Com 9 anos de estudo foi concluída a primeira série do segundo grau. Acima de 12 anos de estudo o indivíduo concluiu com sucesso pelo menos um ano de curso de ensino superior. A primeira série do primeiro grau é feita normalmente com 7 anos de idade (em poucos casos com 6).

Estão presentes na amostra selecionada 200 mil famílias que podem ser classificadas como casais sem filhos, casais com filhos, mulheres solteiras, homens solteiros e mães solteiras.<sup>10</sup> Como o número de pais solteiros era muito pequeno, optou-se por excluir este tipo de família da amostra.<sup>11</sup> A composição da amostra entre os tipos de família, considerando os pesos amostrais, é a seguinte: 15% são casais sem filhos, 62% casais com filhos, as mulheres solteiras representam 5% da amostra, os homens 11% e os 7% restantes são de mães solteiras.

## **2.2. Proxy para Riqueza**

Para inferir o nível de riqueza das famílias, são utilizadas as informações sobre os domicílios referentes a sua estrutura física (material das paredes, da cobertura e número de cômodos), ao seu acesso a serviços urbanos (água, esgoto, luz, coleta de lixo e telefone), à posse de bens duráveis (fogão, geladeira e filtro de água) e ao número de empregados domésticos. É levado em consideração também o fato de a família ser proprietária do imóvel onde reside e / ou proprietária de algum outro imóvel. Além disso, consideram-se também o recebimento de doações de não morador e o recebimento de juros de poupança ou de outras aplicações financeiras.

As características mencionadas acima estão resumidas em 20 variáveis. Fazendo uma combinação linear entre elas, construiu-se uma medida de riqueza. O ideal seria que nela estivesse contida o máximo de informação que se obtém olhando para as 20 variáveis. Foi escolhida, então, aquela combinação linear destas variáveis que melhor expressa a variação amostral observada entre elas. Segundo a análise de componentes principais, a combinação linear que possui a

---

<sup>10</sup> Considerando os pesos amostrais, são, em média, 4 milhões de famílias por ano.

<sup>11</sup> Lembrando que são pais solteiros, com idade entre 20 e 29 anos, que são chefes de um domicílio particular permanente. Só havia 660 pais solteiros com essas características, em todos os anos. A

maior variância, dentre todas as combinações lineares possíveis, é aquela que utiliza como pesos os elementos do autovetor associado ao maior autovalor da matriz de correlação.<sup>12</sup> É preciso enfatizar que a *proxy* para riqueza não se encontra em unidades monetárias. O nível desta variável é irrelevante, sendo importante apenas para efeitos comparativos.

Na tabela 1 estão listados os autovalores da matriz de correlação. Pode-se perceber, a partir desta tabela, que as 20 variáveis podem ser resumidas de forma satisfatória em apenas uma (o primeiro componente principal). Cada autovalor e seu respectivo autovetor permitem o cálculo de um componente principal. Cada componente corresponde a uma combinação linear das 20 variáveis originais, usando os elementos do autovetor como pesos. O autovalor indica o quanto da variância o componente consegue explicar.<sup>13</sup> A terceira coluna da tabela 1 mostra a diferença entre um autovalor e o autovalor imediatamente abaixo. É comumente aceito que uma queda brusca nesse valor, como a que se observa de 3,76 para 0,25, sugere que o autovalor subsequente é só ruído amostral. Então, será utilizado apenas o primeiro componente principal para exprimir o que se observa com as 20 variáveis. De acordo com a coluna de proporção, ele responde por 25% da variância amostral.

[Ver tabela 1 no Apêndice]

A figura 1 apresenta a riqueza média por tipo de família para os anos de 1981 a 2001. Esta figura é composta por 3 gráficos. No primeiro, observa-se de que forma evoluiu a riqueza média dos diferentes tipos de família. Como o nível da variável de riqueza não está em valores monetários, outros 2 gráficos são apresentados. O segundo expressa, a cada ano, a riqueza média das famílias com relação à riqueza média das mulheres solteiras. No terceiro, para cada tipo de família, a riqueza média é apresentada com relação àquela obtida no ano de 1981. O nível de riqueza aumenta ao longo do período analisado para todos os tipos de família. Esse crescimento é maior, entretanto, para as famílias com menor riqueza média, de forma que a riqueza parece convergir. Mães solteiras têm o menor nível de riqueza, mas apresentaram maior taxa de crescimento. A partir de 1997 elas

---

maioria dos pais solteiros na PNAD não mora em um domicílio próprio, morando ainda com os pais.

<sup>12</sup> Optou-se pela matriz de correlação devido às variações nas unidades de medida das variáveis utilizadas.

<sup>13</sup> Para maiores detalhes, ver Johnson e Wichern (1992).

alcançam os casais com filhos e os homens solteiros. A ordenação das famílias por nível de riqueza não sofre alteração considerável. Mulheres solteiras possuem o nível mais elevado, seguidas pelos casais sem filhos. Casais com filhos e homens solteiros têm uma trajetória bem parecida e, como já mencionado, são alcançados mais a frente pelas mães solteiras.

[Ver figura 1 no Apêndice]

### **2.3. Tipos de Família**

Os dados da tabela 2 permitem caracterizar os diferentes tipos de família. Além do nível de riqueza mais elevado, as mulheres solteiras apresentam, em média, maior escolaridade. Entre os casais, percebe-se que a média de anos de estudo dos chefes é bem próxima à média dos respectivos cônjuges. A ordenação dos tipos de família por anos de estudo é a mesma que a feita por nível de riqueza: mulheres solteiras, casais sem filhos, homens solteiros, casais com filhos e mães solteiras. Essas diferenças de anos de estudo e riqueza não parecem refletir diferenças de idade. A idade média não varia muito entre os tipos de família e não é verdade que os mais educados e ricos estão associados à idade maior. Os anos de estudo dos empresários são superiores aos dos não empresários. Essa diferença é maior para mães solteiras (4,61 na média) e para os cônjuges (4,47 para os sem filhos e 4,97 para os com filhos). Empresários também estão associados a maior riqueza e maior idade.

A proporção de empresários é maior entre os chefes dos casais sem filhos. Entre as mulheres solteiras, a proporção de empresários só é maior do que à observada para mães solteiras, grupo com menor nível de riqueza e educação. A proporção de cônjuges empresários é bem inferior à encontrada para os chefes. Quando se calcula a proporção de chefes empresários entre aqueles chefes que possuem cônjuges empresários e a proporção de cônjuges empresários entre aqueles que possuem chefes empresários, parece que possuir um parceiro empresário aumenta as chances de ser também um empresário. O rendimento

médio dos empresários é superior ao dos não empresários, para qualquer tipo de família, tanto para chefes quanto para cônjuges.<sup>14</sup>

A Tabela 2 também fornece alguma informação sobre as crianças. O número médio de filhos é ligeiramente maior para mães solteiras do que para casais com filhos. A proporção de filhos trabalhando é, da mesma forma, maior para esse tipo de família. Por fim, a proporção de filhos com anos de estudo maior ou igual a média de anos de estudo por idade é maior para os casais. Parece, a princípio, que os filhos de mães solteiras estão em pior situação. Ainda com relação às crianças, a figura 2 mostra a distribuição e a distribuição acumulada das idades dos filhos de mães solteiras e de casais. Observa-se que metade dos filhos de casais tem dois anos ou menos de idade, enquanto metade dos filhos de mães solteiras tem idade abaixo de quatro anos, inclusive. Apenas 1% dos filhos de casais tem idade maior que dez anos, enquanto que para mães solteiras esse percentual é de 7%.

Mais uma informação adicional é que grande parte das famílias da amostra pertence ao meio urbano. Apenas 5% das mulheres solteiras residem em áreas rurais. A maior proporção é para casais com filhos: 25% deles moram no meio rural.

[Ver tabela 2 e figura 2 no Apêndice]

#### **2.4. Relação entre Riqueza e Decisões Intra-Familiares**

Antes da análise formal que será apresentada na próxima seção, alguma evidência de que os agentes estão restritos a crédito nas suas escolhas pode ser obtida na figura 3. Nela, os chefes de família estão classificados de acordo com seu tipo e, em cada tipo, por decís de riqueza. Desse modo, pode-se observar como o seu comportamento varia na medida em que vão se tornando cada vez menos dependentes de crédito. Esta figura apresenta em colunas a proporção de empresários em cada tipo de família e decil de riqueza.

[Ver figura 3 no Apêndice]

Com o aumento da riqueza, a restrição de crédito é relaxada e espera-se que um maior número de indivíduos consiga abrir seu próprio negócio. É possível

---

<sup>14</sup> Foram usados os deflatores sugeridos por Corseuil e Foguel (2002), estando os rendimentos em valores de janeiro de 2002.

observar essa tendência na Figura 3. Para todos os tipos de família, a proporção de empresários se eleva quando olhamos em direção aos grupos com um nível de riqueza mais elevado. Observe ainda que esse aumento não é gradual. Para homens solteiros, casais sem filhos e casais com filhos, a proporção de empresários parece subir de patamar em torno do quarto decil e só tornar a se elevar novamente depois do oitavo decil. Isto sugere a existência de custos fixos ou indivisibilidade do investimento.<sup>15</sup> Não adianta a riqueza aumentar um pouco, o aumento deve ser substancial para que a atividade empresarial seja viável. Para as mães solteiras, a proporção de empresários só se altera no último decil, o que indica um custo de entrada mais elevado do que para os demais grupos. A proporção de empresários entre as mulheres solteiras apenas se eleva após o oitavo decil.

As famílias que são suficientemente ricas não estarão restritas nas suas decisões de investimento. Assim, o comportamento dos chefes das famílias que estão entre as 10% mais ricas de sua categoria nos dá uma pista de que comportamento esperar quando crédito não é um problema. Então, a diferença na proporção de empresários entre o primeiro e o décimo decil indica o quão ativa estava a restrição de crédito. A falta de crédito parece ser particularmente importante para casais sem filhos. Entretanto, para as mulheres solteiras e mães solteiras, mesmo no último decil, a proporção de empresários só chega a 2 e 3%, respectivamente. A situação das mulheres solteiras é mais surpreendente dado que são o grupo mais rico e mais bem educado. Note ainda, que sua proporção de empresários não se altera do nono para o último decil. Parece que por razões de outra natureza o desejo de se tornar um empresário é menor nesse tipo de família. As mães solteiras têm menor nível de educação e de riqueza, mas provavelmente possuem maior necessidade de ter uma ocupação mais bem remunerada. No caso das mães solteiras é possível que mesmo as que estão no último decil de riqueza ainda estejam fortemente restritas a crédito.

A figura 3 também apresenta, em linhas no eixo secundário, a riqueza média por decil como proporção da riqueza média do último decil de riqueza das mulheres solteiras. Desse modo, é possível acompanhar os saltos de riqueza entre os decís. Estes, inclusive, vão se tornando cada vez menores ao longo dos decís.

---

<sup>15</sup> Lembre-se que a hipótese de não convexidade tecnológica é fundamental em vários artigos

As curvas de riqueza média praticamente não se cruzam, assim como foi observado no gráfico da riqueza média ao longo dos anos (figura 1). Outra semelhança com o que se observa ao longo dos anos é a ordenação dos tipos de família por riqueza.

A figura 4 apresenta as mesmas informações da figura anterior, mas para os cônjuges. Ela também expressa uma relação positiva entre a proporção de empresários e a riqueza. É possível perceber, pela comparação desta figura com a anterior, que, para os casais, o número de empresários entre os chefes da família é bem superior ao que se observa entre seus respectivos cônjuges. A proporção de empresários entre os cônjuges dos casais com filhos é inferior à encontrada entre qualquer outro membro de qualquer tipo de família.

[Ver figura 4 no Apêndice]

Esse mesmo tipo de análise preliminar sobre a relação entre a riqueza e a decisão de tornar-se um empresário pode ser feito para a decisão dos pais sobre a educação dos filhos. As mães solteiras e os casais com filhos são classificados por decís de riqueza nas figuras 5 e 6.<sup>16</sup> Na figura 5, é apresentada, para cada tipo de família e decil de riqueza, a proporção de crianças que tem um nível educacional igual, ou superior, à média para as crianças da mesma idade.<sup>17</sup> Optou-se por apresentar este dado para meninas e meninos em separado. A relação entre a riqueza e a educação dos filhos não se mostrou tão clara quanto a que se observa entre riqueza e atividade empresarial. Existem outros fatores fundamentais, como características pessoais e dos pais, que afetam o desempenho educacional das crianças. Possivelmente, por ser não condicional a esses fatores, nenhum indício sobre o impacto da riqueza na educação das crianças pode ser obtido na figura 5. É interessante notar, entretanto, um padrão que se repete para cada decil de riqueza: os filhos de mães solteiras são menos educados do que as crianças que moram com ambos os pais e, para o dois tipos de família, as meninas são mais educadas que os meninos.

[Ver figura 5 no Apêndice]

---

teóricos. Em alguns deles, ela está associada à presença de armadilhas de pobreza.

<sup>16</sup> Lembrando que essa classificação é feita para cada tipo de família em isolado. Assim, no primeiro decil estão as mães solteiras mais pobres comparadas com as demais mães solteiras e não comparado com todas as famílias, independente do tipo.

<sup>17</sup> Calculou-se para cada ano e região a média de anos de estudo para cada idade. Em seguida, para cada criança obteve-se a diferença entre seus anos de estudo e essa média.



Na figura 6, é apresentada a proporção de crianças que trabalham. Ao contrário da figura anterior, nesta é possível observar uma tendência. A proporção de crianças trabalhando se reduz do primeiro para o décimo decil. Em quase todos os decís, observa-se uma ordem (que é o inverso da que se observou para educação): os filhos de mães solteiras trabalham mais e os meninos trabalham mais que as meninas.

[Ver figura 6 no Apêndice]

Esta seção apresentou a base de dados, construída a partir das informações contidas nas PNAD's. Destacam-se, na construção da base, a seleção de jovens e a criação da variável de riqueza, que foram pensados de modo a permitir maior clareza na identificação do efeito da riqueza nas decisões de investimento das famílias e que são estratégias mais características deste trabalho. Na última subseção, uma primeira investigação dos dados revelou que entre os indivíduos de maior riqueza a frequência de empresários é maior e o trabalho infantil é mais raro. Isso sugere que a riqueza está afetando a decisão de se tornar um empresário e a decisão a respeito da condição de ocupação das crianças, enquanto não parece ser relevante para a escolaridade dos filhos. A próxima seção apresenta uma análise mais conclusiva destas relações.