

6 Identificação

Para resolver o problema de variáveis omitidas descrito no capítulo anterior, é necessário encontrar uma fonte de variação exógena para o valor recebido do Programa Bolsa Escola. É fato que a idade dos filhos afeta o valor recebido do Programa Bolsa Escola. Famílias com dois filhos com menos de seis anos de idade não podem receber dinheiro do programa. Por outro lado, famílias com dois filhos entre 6 e 15 anos são elegíveis para receber 30 reais do programa.

É difícil, no entanto, argumentar que a idade dos filhos seja exógena às características das famílias. Famílias mais pobres tendem a ter mais filhos e, por isso, tem maior probabilidade de ter filhos de todas as idades. Além disso, a idade dos pais afeta a idade dos filhos. Pode acontecer também da chance de receber dinheiro do programa afetar a decisão, de ter filhos, de famílias pobres.

Imagine, no entanto, dois grupos de famílias com dois filhos. As famílias do primeiro grupo possuem um filho em idade de receber o benefício (entre 6 e 15 anos) e um filho com idade entre 4 e 5 anos. As famílias do segundo grupo possuem dois filhos em idade de receber o benefício. Tem-se ainda que todas as famílias recebem algum benefício. A primeira hipótese de identificação adotada nessa dissertação é que o fato de uma família pertencer ao primeiro ou segundo grupo afeta, de maneira exógena, o valor recebido do programa. Essa hipótese é menos drástica do que adotar que a idade dos filhos afeta o valor recebido do Programa Bolsa Escola de forma exógena. Além disso faz sentido imaginar que a variação é exógena dado que as famílias de ambos os grupos são muito parecidas (são elegíveis, têm dois filhos e recebem o benefício), tendo como principal diferença o fato de um filho não receber o benefício por ter uma idade um pouco abaixo da necessária. Note ainda que filhos de 4 anos de idade teriam nascido, no máximo em 1999 (dado que a pesquisa terminou em junho de 2003), ou seja, antes da criação do Programa Bolsa Escola Federal, de forma que a decisão de ter esses filhos não pode ter sido afetada pela possibilidade de receber o benefício.

A segunda hipótese de identificação adotada nesse trabalho surge de uma idéia análoga a descrita acima. Tome dois grupos de famílias com dois filhos. As famílias do primeiro grupo possuem um filho em idade de receber o benefício e um filho com idade entre 16 e 17 anos. As famílias do segundo grupo possuem dois filhos em idade de receber o benefício. Além disso, todas as famílias recebem algum benefício. A segunda hipótese de identificação assume que o fato de uma família pertencer ao primeiro ou segundo grupo afeta, de maneira exógena, o valor recebido do programa. Nesse caso a idéia é a mesma adotada acima, mas a única diferença entre as famílias é que as famílias do primeiro grupo possuem um filho com idade um pouco acima da idade limite para receber o benefício. Note que todos os filhos, nesse caso, também nasceram antes de 1999.

A terceira hipótese de identificação também assume que o fato de uma família pertencer ao primeiro ou segundo grupo afeta, de maneira exógena, o valor recebido do programa. Nesse caso os grupos tratam de famílias com três filhos. As famílias do primeiro grupo possuem dois filhos em idade de receber o benefício e um filho com idade entre 4 e 5 anos. As famílias do segundo grupo possuem três filhos em idade de receber o benefício. É importante lembrar que todas as famílias recebem algum benefício. Novamente, todos os filhos considerados nesse caso nasceram antes de 1999. A idéia da exogeneidade continua baseada no fato de que as famílias consideradas são muito semelhantes, possuindo apenas uma diferença marginal na idade de seus filhos.

A quarta hipótese de identificação é igual a terceira hipótese de identificação. A única diferença é que o primeiro grupo da quarta hipótese de identificação contém famílias com dois filhos em idade de receber o benefício e um filho com idade entre 16 e 17 anos. Novamente todos os filhos considerados nesse caso nasceram antes de 1999. Além disso, a idéia da exogeneidade permanece a mesma.

Pode-se então condensar todas as hipóteses de identificação em uma. A hipótese de identificação assume que tomando-se duas famílias com o mesmo número de filhos o fato de uma família possuir um filho um pouco abaixo ou um pouco acima da faixa etária agraciada com o benefício é não correlacionado com as características dessa família. Esse fato, no entanto, é obviamente correlacionado com o valor recebido do Programa Bolsa Escola, de forma que representa uma variação exógena neste.

As hipóteses de identificação desenvolvidas acima são utilizadas para criar quatro exercícios de identificação que visam fornecer a relação causal entre o valor recebido do Programa Bolsa Escola e cada categoria de despesa analisada nesse trabalho. O primeiro exercício de identificação consiste em realizar a regressão padrão descrita no capítulo 5, mas considerando uma amostra diferente. A amostra considerada no primeiro exercício de identificação contém dois grupos de famílias com dois filhos. As famílias do primeiro grupo possuem um filho em idade de receber o benefício e um filho com idade entre 4 e 5 anos. As famílias do segundo grupo possuem dois filhos em idade de receber o benefício. Todas as famílias da amostra recebem algum benefício. Note que essa amostra é exatamente a discutida na formulação da primeira hipótese de identificação.

O segundo exercício consiste em realizar a regressão padrão considerando a amostra discutida na formulação da segunda hipótese de identificação. O terceiro exercício de identificação e o quarto também realizam a regressão padrão, mas o terceiro exercício considera a amostra discutida na formulação da terceira hipótese de identificação e o quarto considera a amostra discutida na formulação da quarta hipótese de identificação.

A figura 1 auxilia na explicação do porquê dos coeficientes β_1 encontrados nos exercícios de identificação fornecerem a relação causal entre o valor recebido do Programa Bolsa Escola e cada categoria de despesa. Suponha que o grupo 1 da figura 1 (losangos azuis) represente o primeiro grupo da amostra usada no primeiro exercício de identificação e que o grupo 2 da figura 1 (quadrados rosas) represente o segundo grupo da amostra utilizada neste exercício. Nesse caso, a disposição dos dados indica que o impacto exógeno (primeira hipótese de identificação) que está fazendo com que as famílias do grupo 2 recebam um benefício maior (um filho a mais recebendo o benefício) também está aumentando o gasto médio dessas famílias na categoria de despesa considerada. O coeficiente β_1 é exatamente a inclinação da tendência linear que aparece na figura 1. Essa inclinação é positiva e reflete o fato que famílias que recebem um valor maior do Programa Bolsa Escola (fazem parte do grupo 2) tendem a gastar mais na categoria de despesa considerada. Como o coeficiente β_1 está sendo mensurado, por hipótese, com base em uma variação exógena do valor recebido do Programa Bolsa Escola, ele não contém viés.

A tabela 11 apresenta a estimativa de β_1 obtida nos quatro exercícios de identificação para quase todas as categorias de despesa analisadas na tabela 1. Além disso, a primeira coluna da tabela 11 fornece os coeficientes β_1 estimados na tabela 8. As linhas da tabela 11 deveriam conter todas as categorias de despesa descritas na tabela 1. Foram, no entanto, omitidas duas categorias de despesa dado que o número muito pequeno de observações positivas para a despesa nessas categorias fez com que o método de máxima verossimilhança não convergisse em nenhum dos quatro exercícios de identificação.

A tabela 12 apresenta a estimativa de β_1 obtida nos quatro exercícios de identificação para quase todas as categorias de despesa analisadas na tabela 2. Além disso, a primeira coluna da tabela 12 fornece os coeficientes β_1 estimados na tabela 9. As linhas da tabela 12 deveriam conter todas as categorias de despesa descritas na tabela 2. Foi, no entanto, omitida uma categoria de despesa descrita na tabela 2, dado que o número muito pequeno de observações positivas para a despesa nessa categoria fez com que o método de máxima verossimilhança não convergisse em nenhum dos quatro exercícios de identificação.

A tabela 13 apresenta a estimativa de β_1 obtida nos quatro exercícios de identificação para um conjunto de categorias de despesa analisadas na tabela 3. Além disso, a primeira coluna da tabela 13 fornece os coeficientes β_1 estimados na tabela 10. As linhas da tabela 13 deveriam conter todas as categorias de despesa contidas na tabela 3. Foram, no entanto, omitidas seis categorias de despesa descritas na tabela 3, dado que o número muito pequeno de observações positivas para a despesa nessas categorias fez com que o método de máxima verossimilhança não convergisse em nenhum dos quatro exercícios de identificação.

Uma observação importante é que o trabalho não tem nada a dizer a respeito do efeito do Programa Bolsa Escola sobre as despesas nas categorias para as quais o método de máxima verossimilhança não convergiu em nenhum dos quatro exercícios de identificação.

Uma das categorias de despesa contidas na tabela 11 é a categoria de despesa total com saúde. Na primeira coluna aparece um coeficiente não significativo igual a 0,011. Esse coeficiente é a estimativa para β_1 apresentada na tabela 9. O primeiro exercício de identificação reporta um coeficiente significativo igual a

0,403. Esse coeficiente é a estimativa de β_1 da regressão padrão quando considera-se a amostra do primeiro exercício de identificação. O segundo exercício de identificação reporta um coeficiente significativo igual a 0,503. Esse coeficiente é o β_1 da regressão padrão que considera a amostra do segundo exercício de identificação. O terceiro exercício de identificação reporta um coeficiente significativo igual a -0,901. Por fim, o quarto exercício de identificação fornece um coeficiente não significativo igual a -0,438.

A análise acima demonstra grande instabilidade do coeficiente β_1 , quando a variável dependente é a categoria de despesa total com saúde. Esse coeficiente é em determinado exercício positivo, significativo e igual a 0,503 e em outro exercício negativo, significativo e igual a -0,901. A grande instabilidade de β_1 é decorrência do número reduzido de famílias da amostra considerada que têm despesa positiva com saúde. Em todos os exercícios de identificação menos de 15% das famílias da amostra têm despesa positiva em saúde.

O coeficiente β_1 é mais estável, ou robusto, em categorias em que uma grande porcentagem das famílias têm despesa positiva. Para garantir que os resultados apresentados nessa tese são robustos, optou-se por considerar uma determinada categoria de despesa, apenas se em todos os exercícios de identificação mais de 50% das famílias têm despesa positiva nessa categoria. Essa regra recebeu o nome de regra de robustez. Para simplificar a análise dos resultados as categorias que atendem a regra de robustez tiveram seus nomes pintados de vermelho nas tabelas 11, 12 e 13. Por exemplo, o nome da categoria de despesa total com roupas, que aparece na tabela 11, está pintado de vermelho, então essa categoria atende a regra de robustez.

Uma investigação geral dos coeficientes das categorias que atendem a regra de robustez indica que esses coeficientes também são bastante instáveis. Essa instabilidade é decorrente do número reduzido de famílias que são consideradas nos exercícios de identificação. Mesmo assim, esses coeficientes são mais estáveis do que os coeficientes encontrados em categorias que não atendem a regra de robustez. Um dado que reflete a maior estabilidade dos coeficientes das categorias que atendem a regra de robustez é que em nenhuma dessas categorias acontece do coeficiente ficar significativo e positivo em um exercício de identificação e significativo e negativo em outro.

Diante de tanta instabilidade é reconfortante encontrar duas categorias de despesa, que atendem a regra de robustez, para as quais a estimativa do coeficiente β_1 é relativamente estável. Essas duas categorias são a de despesa com cereais, leguminosas e hortaliças tuberosas e a de despesa com açúcares e produtos de confeitaria. Os coeficientes estimados nos quatro exercícios de identificação em que a variável dependente é a despesa com cereais, leguminosas e hortaliças tuberosas são positivos. Isso também acontece com os coeficientes estimados nos quatro exercícios de identificação em que a variável dependente é a despesa com açúcares e produtos de confeitaria. Nos dois casos, tem-se que três dos quatro coeficientes são significativos.

No caso em que a variável dependente é a despesa com cereais, leguminosas e hortaliças tuberosas os coeficientes significativos são ou maiores ou parecidos com o coeficiente encontrado na tabela 9. Essa evidência indica que o coeficiente causal de um aumento de um real na bolsa fornecida pelo Programa Bolsa Escola é ou maior ou parecido com o estimado na tabela 9. O presente trabalho adota, no entanto, uma postura conservadora e assume que o efeito causal é o do coeficiente 0,237 apresentado na tabela 9. Dessa forma, o trabalho assume que um aumento de um real na bolsa do programa levaria a um aumento de, aproximadamente, 24 centavos na despesa com cereais, leguminosas e hortaliças tuberosas. Note que os principais produtos contidos nessa categoria de despesa são arroz, feijão e batata (tabela 2a), de forma que esse coeficiente indica que as famílias que recebem a bolsa gastam mais nesses produtos. Por último, é importante lembrar que esse coeficiente foi obtido através de uma regressão tobit, de forma que não representa o efeito parcial médio de uma alteração de um real na bolsa. O cálculo do efeito parcial médio, no entanto, levaria a um coeficiente menor, porém ainda positivo e significativo, não alterando em grande medida o resultado apresentado nesse parágrafo.

No caso em que a variável dependente é a despesa com açúcares e produtos de confeitaria todos os coeficientes significativos são maiores do que o coeficiente estimado na tabela 9. Essa evidência indica que o coeficiente causal de um aumento de um real na bolsa fornecida pelo Programa Bolsa Escola é maior do que o estimado na tabela 9. O presente trabalho adota novamente uma postura conservadora e, mesmo diante da evidência de que o efeito causal é maior do que o estimado na tabela 9, assume que o efeito causal é o do coeficiente 0,108

apresentado na tabela 9. Dessa forma, o trabalho assume que um aumento de um real na bolsa do programa levaria a um aumento de, aproximadamente, 11 centavos na despesa com açúcares e produtos de confeitaria. É importante lembrar que esse coeficiente foi obtido em uma regressão tobit, de forma que não representa o efeito parcial médio. O cálculo do efeito parcial médio levaria a um coeficiente menor, mas esse coeficiente continuaria sendo positivo e significativo, não alterando em grande medida o resultado apresentado nesse parágrafo. Pode-se pensar na categoria de despesa com açúcares e produtos de confeitaria como sendo a categoria de despesa com doces.

Outras duas categorias, que atendem a regra de robustez, que também apresentam certa estabilidade em seus coeficientes são a de despesa com bebidas não alcoólicas e infusões e a de despesa com artigos de limpeza doméstica. Nesses dois casos todos os coeficientes estimados nos exercícios de identificação são positivos e dois deles são significativos. Novamente opta-se pelo conservadorismo, de forma que os coeficientes considerados são os que aparecem na tabela 9. Apesar desses coeficientes serem relativamente estáveis, a evidência mais forte é a do efeito do Programa Bolsa Escola sobre as despesas com cereais, leguminosas e hortaliças tuberosas e sobre as despesas com açúcares e produtos de confeitaria, por este motivo, os resultados apresentados só consideram que exista evidência desses dois efeitos. Note que a conclusão apresentada abaixo não se alteraria muito ao considerar-se também os efeitos sobre a despesa com bebidas não alcoólicas e infusões e sobre a despesa com artigos de limpeza doméstica.

Os coeficientes das outras categorias que atendem a regra de robustez não parecem ser robustos o suficiente para serem confiáveis. Esses coeficientes ficam significativos em poucos casos e, são em sua maioria instáveis. Dessa forma, parece não haver efeito do Programa Bolsa Escola nessas categorias de despesa.

Os resultados apresentados nas tabelas 11, 12 e 13 indicam que o Programa Bolsa Escola afeta a despesa em duas categorias de alimentação (despesa em cereais leguminosas e hortaliças tuberosas e despesa com açúcares e produtos de confeitaria) e que não há evidências de que esse programa afete nenhuma outra categoria de despesa desenvolvida nessa dissertação. Então, como foi explicado na introdução, o Programa Bolsa Escola afeta a despesa em alimentação das famílias e não afeta a despesa em educação, dado que não afeta a despesa em nenhuma categoria presente na tabela 3.

Existe uma grande literatura que discute se o Programa Bolsa Escola Federal aumenta a frequência escolar, mas essa literatura não será discutida nessa dissertação. Esse trabalho apenas demonstra que independentemente de se o programa afeta ou não a frequência, existe um benefício claro do programa, que é o aumento da despesa com alimentação, que pode impactar no rendimento escolar (Grantham-McGregor et al, 1991) e na saúde das crianças das famílias agraciadas com o benefício.

Hoddinott e Skoufias (2004) encontram que o Progesa, o Programa Bolsa Escola do México, gera um impacto positivo sobre as despesas das famílias em alimentos. Desse ponto de vista o resultado encontrado nesse trabalho é análogo ao de Hoddinott e Skoufias (2004). A diferença é que enquanto, nesse trabalho, o aumento de despesas ocorre em cereais, leguminosas e hortaliças tuberosas e em açúcares e produtos de confeitaria (arroz, feijão, batata e doces), no trabalho de Hoddinott e Skoufias (2004) o aumento se concentra em frutas, vegetais e produtos de origem animal. Esses autores argumentam que a existência de *platicas*, palestras compulsórias para as mães de famílias agraciadas com o benefício em que são oferecidas aulas de nutrição, pode ser uma justificativa para que o efeito tenha se concentrado em tipos de despesa que representam uma melhoria não só na quantidade, mas na qualidade da alimentação. No Programa Bolsa Escola não existe um mecanismo semelhante às *platicas*. Isso pode explicar o fato do programa gerar um aumento da quantidade consumida, mas sem melhorar a qualidade, dado que o aumento se concentra em compras de arroz, feijão, batatas e doces.

Goldberger (1989) desenvolve um modelo cujos resultados implicam que um aumento do gasto (ou das transferências) do governo em educação pode aumentar, diminuir ou não afetar o gasto das famílias em educação. Esse resultado gera uma dúvida a respeito da eficácia de programas em que o governo aumenta a sua despesa (ou as suas transferências) em educação. Dessa forma, os benefícios do Programa Bolsa Escola poderiam ser mais do que compensados caso houvesse uma queda no gasto em educação das famílias. O fato dos resultados apresentados indicarem que não há efeito do Programa Bolsa Escola sobre as despesas das famílias em educação indica que, no caso desse programa, não existe o risco de que o aumento das transferências esteja gerando uma queda nas despesas em educação.

A próxima seção conclui.