

# Les choix d'activité des femmes en couple : une approche longitudinale

Stéfan Lollivier\*

---

L'analyse de l'activité des femmes en couple repose généralement sur une hypothèse de spécialisation qui s'opérerait au sein du couple. L'homme est supposé disposer d'une meilleure rentabilité sur le marché du travail et la femme est supposée prendre prioritairement en charge les activités domestiques du foyer. Dans ce cadre, les travaux économétriques ont montré la complexité de la décision d'activité des femmes en analysant le lien entre fécondité et participation au marché du travail, les effets incitatifs ou désincitatifs de la politique fiscale ou le coût lié aux jeunes enfants. Ces résultats présentent l'inconvénient majeur de ne pas tenir compte de l'hétérogénéité des comportements individuels dus aux aptitudes ou aux préférences personnelles des individus et à la façon dont celles-ci s'inscrivent dans leurs trajectoires. L'exploitation du volet français du *Panel européen des ménages* et l'utilisation de nouvelles méthodes économétriques permettent d'affiner l'analyse du comportement d'activité des femmes vivant en couple sur des données longitudinales.

Cette nouvelle approche remet en cause l'influence habituellement admise de certains déterminants du comportement d'activité des femmes. Si, comme attendu, les femmes sont d'autant plus actives qu'elles sont diplômées, le revenu du conjoint joue peu sur la décision d'activité, au moins à court terme. C'est davantage son niveau moyen ou sur longue période qui importe. De même, si le jeune enfant pèse temporairement sur l'activité des femmes, le nombre total d'enfants semble donner lieu à un arbitrage entre fécondité et activité plutôt sur l'ensemble du cycle de vie, même si les coûts induits par la charge des enfants comptent également dans la décision d'activité de nombreuses femmes.

---

\* Stéfan Lollivier est directeur de l'Ensaë.  
Les noms et dates entre parenthèses renvoient à la bibliographie en fin d'article.

L'activité des femmes vivant en couple a sans doute été l'un des sujets les plus étudiés dans la littérature économique et économétrique. L'une des raisons en est probablement que le comportement de choix d'activité illustre de façon quasi immédiate l'un des aspects de la théorie du consommateur : l'arbitrage entre travail et loisir. Selon celui-ci, l'individu rationnel arbitre l'allocation de son temps entre activité rémunérée permettant d'acquérir des biens de consommation et loisir, ces deux derniers lui étant nécessaires pour obtenir un certain niveau d'utilité. D'après ce schéma, le comportement est optimal lorsque le surcroît d'utilité obtenu grâce aux biens de consommation permis par la dernière heure travaillée est égal à l'accroissement de bien-être que lui procurerait la même heure passée en temps de loisir. La réalité s'éloigne cependant de ce comportement stylisé. En premier lieu, l'individu n'est pas totalement maître de sa quantité d'offre de travail. En France, notamment, la durée de travail relève de la loi ou de conventions collectives et présente un caractère indivisible entre temps complet, temps partiel ou inactivité. Mais le travail en temps partiel est lui-même soumis à des contraintes réglementaires. En outre, la décision de travailler à temps partiel renseigne mal sur l'offre de travail de l'individu puisqu'une fraction importante des personnes concernées occupe un emploi à temps partiel dit contraint, c'est-à-dire imposé par l'employeur. Au total, la décision d'activité porte davantage sur le fait de travailler ou non que sur la durée travaillée elle-même. C'est pourquoi un bon nombre d'études empiriques s'intéresse à la décision d'activité plutôt qu'au nombre d'heures travaillées. Du fait de cette indivisibilité et de la nécessité de disposer de ressources pour acheter des biens de consommation, la décision d'activité ne se pose même pas lorsque le ménage ne comporte qu'un seul adulte en âge (ou en état) de travailler.

Un comportement d'offre de travail peut, en revanche, se manifester lorsque le ménage est composé d'un couple, avec ou sans enfants. L'activité d'un seul adulte peut apporter des ressources monétaires jugées suffisantes pour que l'autre ne se porte pas sur le marché du travail. Même si la question se pose en des termes symétriques pour les deux membres du couple, c'est dans la quasi-totalité des cas sur la femme que va se reporter la décision d'activité. Dans une étude un peu ancienne, Becker (1983) explique que la femme reste au foyer prioritairement à l'homme à l'issue d'un processus de spécialisation, l'homme, plus diplômé, disposant d'une meilleure rentabilité sur le marché du travail.

Le couple a donc collectivement intérêt à ce que la femme se spécialise dans les tâches ménagères, libérant le temps de son conjoint, à charge pour lui de le valoriser sur le marché du travail. Même si ce fait demeure aujourd'hui, l'explication n'est plus qu'à moitié convaincante, les femmes ayant, notamment pour les générations récentes, des diplômes au moins équivalents à ceux des hommes.

Outre l'intérêt théorique, la seconde raison pour laquelle les économistes se sont penchés sur l'activité des femmes en couple est la nécessité de disposer de projections à moyen terme de la population active. La connaissance des déterminants structurels et conjoncturels de la flexion des taux d'activité féminins est alors précieuse pour déterminer le nombre d'actifs dans le futur, même si pratiquement, on se limite souvent à projeter les tendances antérieures.

### **Une approche jusqu'ici essentiellement transversale**

De nombreux travaux portant sur les déterminants de l'activité des femmes en couple ont alimenté la littérature économique depuis les années 70. Dans le sillage de Mincer (1963), bon nombre d'entre eux ont cherché à expliciter les liens entre fécondité et participation au marché du travail. Ces travaux ont, en outre, été l'occasion d'avancées économétriques importantes, notamment dans la façon de traiter le fait que la sous-population des femmes actives n'est pas représentative de l'ensemble (biais de « sélectivité », Heckman, 1974). En France, des travaux analogues sont apparus plus tardivement. Ils ont visé soit à décrire la situation des femmes en couple (Riboud, 1985 ; Lollivier, 1984), soit à évaluer les aspects incitatifs ou désincitatifs de la politique fiscale (Bourguignon, 1986 ; Dagsvik, Laisney, Strom et Osterwald, 1988). Dans les années 90, les développements ont approfondi la description des comportements. Eggink, Hop et Van Praag (1994) ont ainsi considéré l'offre conjointe de travail au sein du ménage, en prenant en compte le rôle de l'épouse et de son conjoint : le comportement de la femme est ainsi plus sensible aux variations de l'environnement socio-économique que celui de l'homme. Cette approche a été approfondie par Chiappori, Fortin et Lacroix (1998), puis par Moreau (2000), qui ont mis en évidence, sur des données individuelles, une rationalité collective des choix d'activité au sein du couple. Pour sa part, Van Der Klaauw (1996) a considéré que l'horizon féminin s'étendait au cycle de vie, et décrit un modèle intertemporel

de choix simultané d'offre de travail et de statut matrimonial : les gains liés au mariage seraient d'autant plus faibles que les salaires potentiels de l'épouse seraient élevés. Ne pas prendre en compte l'endogénéité des choix matrimoniaux surestimerait, en particulier, le rôle joué par le salaire du mari. Enfin, différents travaux ont cherché à mieux décrire les coûts de participation au marché du travail. Lanot et Robin (1997) relèvent ainsi des coûts élevés liés aux jeunes enfants, expliquant la moindre participation des femmes ayant des enfants de moins de trois ans.

### **L'apport de l'approche longitudinale : la prise en compte de l'hétérogénéité individuelle**

Parallèlement, et essentiellement outre-atlantique, la réflexion sur l'estimation des comportements d'activité féminins s'est poursuivie avec l'utilisation de données de panel. Dès 1984, Chamberlain montrait comment les développements portant sur l'hétérogénéité des comportements pouvaient s'appliquer à ce type de données et conduire à une remise en question des résultats obtenus en coupe transversale. Ces travaux ont débouché sur des analyses de plus en plus sophistiquées s'intéressant à la spécification des modèles en présence d'hétérogénéité, ainsi qu'à leur estimation (Keane, 1994 ; Arrelano et Carrasco, 1996). On sait pourtant que le simple examen des comportements en coupe transversale peut conduire à des erreurs d'interprétation. Pour que les résultats obtenus en coupe soient valides, il faut supposer que les consommateurs sont tous identiques. Si leurs préférences diffèrent, on ne sait plus si les disparités constatées une année donnée sont porteuses d'une logique économique ou simplement révélatrices de différences de goûts ou de capacités. Un bon exemple de corrélation trompeuse en coupe instantanée est le lien que l'on observe entre salaire plus élevé et utilisation de nouvelles technologies : les salariés qui utilisent dans l'entreprise des technologies innovantes disposent de meilleurs salaires. Pourtant, l'examen de données longitudinales montre que fréquemment ces salariés disposaient déjà de salaires plus élevés avant l'introduction des nouvelles technologies. Une sélection s'est, en fait, simplement produite au sein de l'entreprise, mettant en évidence une disparité latente. La nouvelle technologie n'a fait qu'amplifier la tendance antérieure.

Cependant, peu de travaux, en France, se sont intéressés à l'étude des comportements d'activité en ayant recours à la dimension longitudi-

nale, sans doute faute de données adéquates. Les travaux sur panel se sont plutôt intéressés à des modélisations utilisant des variables quantitatives (salaires, investissement, demande de travail, etc.) techniquement plus aisées à utiliser. Les rares études sur les comportements représentés par des variables qualitatives ont pourtant montré l'ampleur de l'hétérogénéité individuelle non observable (Magnac, 1997). En utilisant le volet français du *Panel européen des ménages*, on réestime ici, sur données longitudinales, les comportements d'activité des femmes vivant en couple. Par rapport aux travaux antérieurs (Lollivier, 1995), l'originalité de l'approche consiste à prendre en compte une éventuelle corrélation entre les effets individuels et les variables explicatives. Ceci permet notamment de mieux expliciter le rôle des variables explicatives observé en coupe. En effet, une année donnée, il est impossible de discerner si l'impact apparent d'une variable qui, comme l'âge ou le revenu, se modifie au cours du temps correspond à son influence véritable. Avoir un certain âge correspond à la fois à une étape du cycle de vie et au fait d'être né une année donnée. Disposer d'un certain revenu peut ne pas entraîner les mêmes effets selon que ce revenu est le revenu habituellement perçu ou un revenu correspondant à une évolution particulière. Les quatre premières vagues du *Panel* autorisent un suivi annuel de 1994 à 1997 qui permet de faire la part entre le rôle effectif joué par les variables et le fait qu'elles sont des révélateurs d'une hétérogénéité individuelle inobservée et constante dans le temps (cf. encadré 1). Parmi les femmes vivant en couple, on se restreint aux femmes nées entre 1938 et 1973, âgées donc de 21 à 56 ans en 1994, et dont les conjoints sont salariés ou anciens salariés. La restriction sur l'âge permet de se limiter aux périodes habituelles de vie active. Celle sur le statut du conjoint est classique dans ce genre d'analyse. Elle permet d'exclure les femmes d'indépendants dont le statut éventuel d'aide familial est de nature différente de la notion habituelle d'activité chez les salariés. En outre, le revenu des indépendants n'est pas toujours comparable à celui des salariés, en raison des règles fiscales et comptables en vigueur en France. Deux autres restrictions ont été introduites : la femme doit avoir conservé le même conjoint au cours de la période et ce dernier rester salarié ou ancien salarié. Mais ces restrictions ne concernent qu'un nombre extrêmement réduit de femmes (1). En dernier lieu, le *Panel* n'a pas été

1. Seules 53 femmes dans cette situation ont été retirées de l'échantillon.

cylindré, c'est-à-dire qu'on a conservé dans le champ toutes les femmes éligibles, même celles qui ont refusé de répondre à l'enquête temporairement ou durablement. Bien qu'il complique les calculs, ce choix permet de ne pas trop restreindre l'échantillon et surtout de lui conserver un maximum de représentativité.

### Plus les femmes sont diplômées, plus elles sont actives

Au total, l'échantillon se compose, en 1994, de 2 913 femmes (cf. tableau 1). L'attrition est limitée, car on retrouve respectivement 90 %, 85 % et 74 % des femmes en 1995, 1996 et

#### Encadré 1

### LE PANEL EUROPÉEN DES MÉNAGES : UN ÉCHANTILLON LONGITUDINAL

Le *Panel européen des ménages* est une opération communautaire, coordonnée par Eurostat, dont l'objectif principal est de suivre les dynamiques d'emploi et de revenu.

La première vague a eu lieu en 1994 dans les douze États formant l'Union européenne de l'époque. L'échantillon total comprenait environ 60 500 ménages sélectionnés de façon aléatoire avec un taux de sondage uniforme. À la fin de 1997, quatre interrogations avaient eu lieu sur les six projetées. Cette enquête permet notamment de fournir des statistiques entre pays afin d'éclairer les aspects sociaux de la politique communautaire.

#### Situation sur le marché du travail et environnement familial

En France, l'Insee assure la gestion et la collecte de l'information. La partie du questionnaire utilisée dans l'article est celle relative aux individus qui vivent dans les ménages interrogés. On utilise les quatre vagues d'enquête de 1994 à 1997, pour lesquelles on dispose d'un fichier homogène. Pour chaque vague, le questionnaire décrit la situation des individus sur le marché du travail, mais aussi leur environnement familial. En outre, la première vague de 1994 recense tous les enfants des individus enquêtés, notamment ceux qui vivent hors du ménage. Le suivi longitudinal permet de mettre à jour le calendrier biographique : naissance des enfants, départ du foyer parental, etc.

Tableau 1  
Caractéristiques de l'échantillon

	En 1944	En 1995	En 1996	En 1997
<b>Âge</b>				
Moins de 25 ans	184	99	53	18
De 25 à 29 ans	400	337	291	220
De 30 à 34 ans	479	433	384	309
De 35 à 39 ans	480	448	435	389
De 40 à 44 ans	475	423	403	362
De 45 à 49 ans	477	438	414	355
De 50 à 54 ans	300	283	303	300
55 ans et plus	118	156	195	204
<b>Diplôme</b>				
Pas de diplôme	538	449	424	368
CEP	402	365	340	290
CAP-BEP	484	436	412	372
BEPC	551	502	470	402
Baccalauréat technique	143	128	118	99
Baccalauréat général	235	215	209	184
Supérieur au baccalauréat	560	522	505	442
<b>Salaire mensuel du conjoint</b> (en francs 1997) (1)				
Conjoint chômeur	202	147	150	126
Moins de 5 800 (884)	209	217	185	143
De 5 800 à 7 100 (884 à 1 082)	317	286	323	258
De 7 100 à 9 500 (1 082 à 1 448)	717	630	522	471
De 9 500 à 13 300 (1 448 à 2 028)	635	585	594	536
De 13 300 à 19 800 (2 028 à 3 018)	412	386	368	334
Plus de 19 800 (3 018)	268	249	241	201
Non déclaré	153	117	95	88
<b>Nombre d'enfants</b>				
Aucun	399	284	222	155
Un	611	563	516	425
Deux	1 071	973	957	867
Trois et plus	832	797	783	710
<b>Ensemble</b>	<b>2 913</b>	<b>2 617</b>	<b>2 478</b>	<b>2 157</b>

1. Les équivalents en euros figurent entre parenthèses.

Lecture : ce tableau indique le nombre de femmes dans chacune des situations.  
Source : Panel européen 1993-1997, Insee.

1997. Le taux d'activité dans l'échantillon est de 73 % (cf. tableau 2). Il s'érode légèrement au cours de la période, essentiellement en raison de la naissance d'enfants, qui, on le verra, pèse sur le comportement d'activité. Entre 30 et 50 ans, âges au cours desquels la mise en couple est pratiquement achevée sans que la mortalité soit encore importante, la proportion des femmes reste quasiment stable. Les femmes plus âgées sont, pour leur part, moins fréquemment actives. Le suivi des générations au cours des quatre années disponibles dans le *Panel* ne fait pas apparaître de dérive importante du comportement d'activité en fonction de la date de naissance (cf. graphique I). Des travaux antérieurs montraient déjà que l'essentiel des particularismes générationnels en matière d'activité était le fait des personnes nées dans l'avant-guerre (Lollivier, 1995). L'activité est d'autant plus fréquente que la femme dispose d'un diplôme de niveau élevé. Ce résultat attendu s'explique dans le cadre du modèle habituel d'arbitrage entre travail et loisir : l'activité est d'autant plus fréquente que les revenus qu'elle procure sont élevés et que les coûts liés à l'activité sont faibles. Le diplôme, qui permet d'accéder à des

emplois mieux rémunérés, joue donc positivement sur cet arbitrage. En revanche, l'influence du revenu du conjoint paraît plus surprenante. À ce niveau de l'analyse, on s'attendrait à ce qu'un revenu du conjoint plus élevé décourage l'activité du fait de la progressivité du système fiscal : les prélèvements sur un même salaire féminin sont d'autant plus élevés que le salaire du mari est fort, puisque la tranche marginale d'imposition est plus grande. Ceci paraît contradictoire avec une augmentation apparente du taux d'activité des femmes avec les revenus des conjoints, lorsque ces derniers sont dans le bas de l'échelle. On verra ultérieurement que ceci provient de la forte homogamie sociale qui conduit les femmes diplômées, donc plus souvent actives, à vivre avec des conjoints eux-mêmes diplômés.

L'une des particularités de l'échantillon est de pouvoir disposer du nombre total d'enfants de la femme et non des seuls enfants présents au domicile du foyer parental. Dans l'échantillon, le nombre d'enfants s'accroît avec l'ancienneté de la génération pour se stabiliser à partir des générations nées avant les années 50

Tableau 2  
Taux d'activité des conjoints de salariés de 1994 à 1997

	En 1944	En 1995	En 1996	En 1997
En %				
<b>Âge</b>				
Moins de 25 ans	76,1	78,8	79,2	77,8
De 25 à 29 ans	81,3	78,3	78,4	73,2
De 30 à 34 ans	74,3	75,1	76,3	75,1
De 35 à 39 ans	75,6	76,6	73,8	76,3
De 40 à 44 ans	74,5	74,0	76,9	79,0
De 45 à 49 ans	71,5	72,4	75,1	75,2
De 50 à 54 ans	63,0	67,8	65,7	68,0
55 ans et plus	51,7	50,6	47,7	47,1
<b>Diplôme</b>				
Pas de diplôme	54,3	51,0	50,5	49,5
CEP	66,4	64,1	62,9	64,5
CAP-BEP	77,1	76,6	76,0	72,8
BEPC	75,0	75,9	76,4	76,4
Baccalauréat technique	88,1	87,5	85,6	84,8
Baccalauréat général	74,0	75,8	74,2	74,5
Supérieur au baccalauréat	86,4	87,7	87,3	88,0
<b>Salaire mensuel du conjoint</b> (en francs 1997) (1)				
Conjoint chômeur	77,7	68,0	66,7	68,3
Moins de 5 800 (884)	68,4	70,5	67,0	65,0
De 5 800 à 7 100 (884 à 1 082)	70,3	69,2	68,7	72,5
De 7 100 à 9 500 (1 082 à 1 448)	72,4	75,9	77,0	76,9
De 9 500 à 13 300 (1 448 à 2 028)	77,6	75,2	77,3	75,6
De 13 300 à 19 800 (2 028 à 3 018)	75,2	78,8	76,6	74,0
Plus de 19 800 (3 018)	66,8	63,1	62,2	61,7
Non déclaré	68,6	69,2	61,1	60,2
<b>Nombre d'enfants</b>				
Aucun	87,2	88,0	87,4	87,7
Un	84,3	86,3	86,8	86,4
Deux	78,2	78,1	78,3	78,2
Trois et plus	51,6	52,1	51,9	53,0
<b>Ensemble</b>	<b>73,1</b>	<b>73,0</b>	<b>72,5</b>	<b>72,2</b>

1. Les équivalents en euros figurent entre parenthèses.

Source : Panel européen 1993-1997, Insee.

(cf. graphique II). La proportion des enfants vivant hors du ménage augmente pour les générations nées avant 1950 et devient prépondérante pour les générations nées avant la fin de la guerre. Un nombre d'enfants élevé va logiquement de pair avec une activité moindre, ce qui est qualitativement conforme avec l'existence de coûts induits par la présence d'enfants (coûts de transport, de garde, etc.).

### La présence d'un jeune enfant modère l'activité

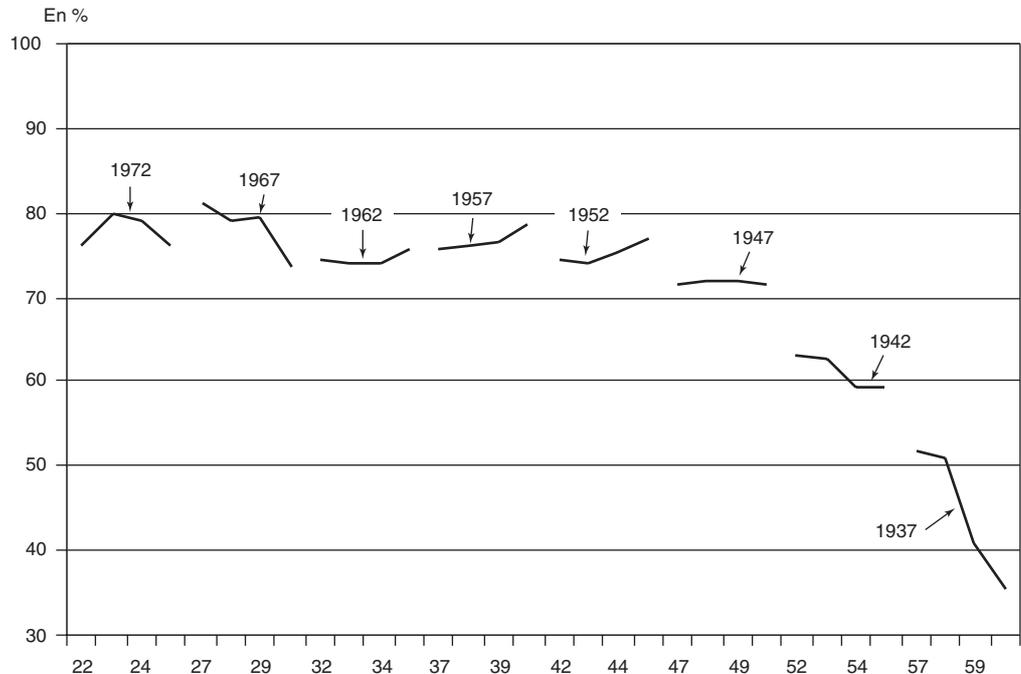
En recourant à une analyse toutes choses égales par ailleurs, simplement en empilant les quatre vagues d'enquête disponibles (cf. *infra* tableau 6, modèle 1 et tableau 7, modèle 1), ces premiers résultats se précisent. Les disparités en matière d'activité sont, à nouveau, particulièrement accentuées en fonction du diplôme : plus le diplôme est élevé, plus l'activité augmente (2). Cependant, les femmes disposant d'un baccalauréat technique sont significativement plus actives que celles ayant un baccalauréat général. Pour sa part, le nombre d'enfants pénalise l'activité, même si les femmes sans enfant et les femmes avec un enfant ont un comportement d'activité

analogue. Les femmes dont le plus jeune enfant a moins de quatre ans sont plus fréquemment absentes du marché du travail. En outre, les femmes dont les enfants ont totalement ou partiellement quitté le foyer parental sont, elles aussi, moins fréquemment actives. Ces trois variables concernant les enfants (nombre, âge et présence dans le foyer parental) apparaissent dans l'estimation de façon additive, c'est-à-dire, par exemple, que le jeune enfant jouera le même rôle quelque soit son rang, comme sa présence ou son absence dans le foyer. Des variantes de cette estimation avec une interactivité plus forte montrent que même si certaines d'entre elles améliorent à la marge le pouvoir explicatif du modèle, elles ne remettent pas en question les effets : le jeune enfant joue seulement un peu moins lorsqu'il est le premier. Pour des besoins de commodité d'estimation et de présentation, on se restreint ici à la version purement additive.

Cette prise en compte du nombre d'enfants dans le modèle fait cependant apparaître un profil d'activité selon l'âge assez différent du profil

2. Faute de disposer d'une méthode simple pour traiter l'endogénéité d'une variable polytomique, on admettra l'exogénéité du diplôme.

Graphique I  
Taux d'activité des conjoints de salariés selon l'âge et la génération



Lecture : le graphique indique pour chaque génération repérée par son année de naissance le taux d'activité à différents âges. Les différentes générations ont été regroupées par classes quinquennales (1957 recouvre ainsi les générations allant de 1955 à 1959). Comme le Panel comporte quatre points, chaque génération est représentée par quatre points qui correspondent à quatre âges consécutifs.  
Source : Panel européen 1993-1997, Insee.

apparent. Au lieu d'être relativement plat jusqu'à 50 ans, ce profil est d'abord croissant puis décroissant, probablement parce que les femmes les plus jeunes ont encore peu d'enfants. Parmi elles, on ne distingue pas entre celles qui vont effectivement avoir peu d'enfants et sans doute des taux d'activité plus élevés et celles qui auront davantage d'enfants et peut-être des taux d'activité plus faibles. En les comparant à la situation de référence de femmes qui ont un seul enfant, donc parmi les plus actives, leur comportement d'activité « moyen » est logiquement plus faible. Pour leur part, les effets de génération sont modestes. Plusieurs variantes montrent que le seul effet significatif concerne les femmes des générations les plus anciennes, un peu moins actives toutes choses égales par ailleurs, notamment au même âge.

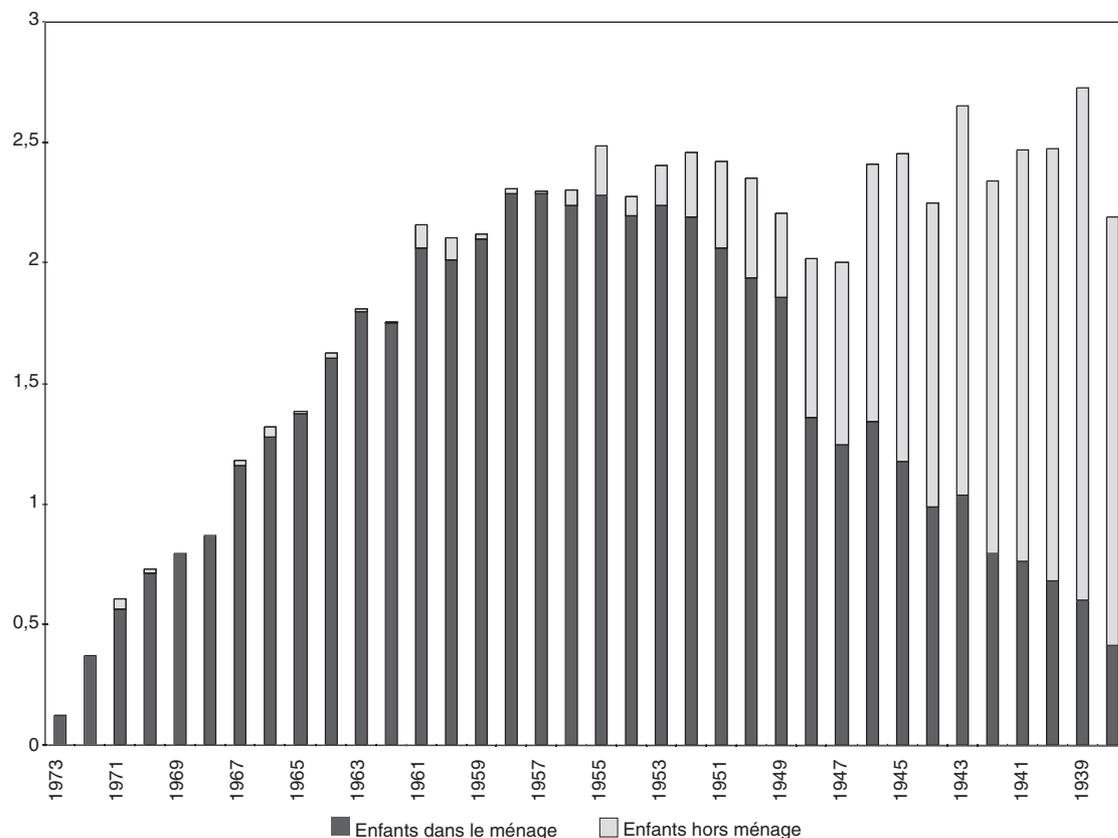
Le dernier intérêt du modèle consiste à faire apparaître un profil d'activité selon le revenu du conjoint plus conforme *a priori* avec les attendus théoriques. À diplôme de la femme donné,

le salaire du conjoint ne joue pas jusqu'à 1 448 euros (9 500 francs) par mois, soit *grosso modo* la médiane de la distribution des salaires de ces conjoints. L'effet négatif est après ce seuil de plus en plus net. Ceci est qualitativement conforme à ce que l'on peut attendre du système fiscal. Celui-ci est neutre pour une moitié des ménages qui ne paient pas l'impôt sur le revenu. Au-delà, les taux marginaux de prélèvements sont de plus en plus élevés, ce qui est cohérent avec un découragement croissant.

### À peu près autant de sorties que d'entrées en activité

Ces résultats recouvrent la plupart de ceux obtenus jusqu'ici, même si la population des conjoints de salariés est toujours considérée comme un ensemble homogène en termes de préférences. Disposer d'un panel permet de discuter cette hypothèse. En premier lieu, on vérifie que les variables utilisées dans l'analyse présentent une variabilité temporelle. Ainsi, les femmes

Graphique II  
Répartition du nombre moyen d'enfants dans les couples de salariés selon l'année de naissance de la mère et leur appartenance au foyer



Source : Panel européen 1993-1997, Insee.

effectuent des transitions entre l'activité et l'inactivité (3). Parmi les femmes inactives une année donnée, un peu moins de 8 % seront actives l'année suivante. À l'inverse, parmi les femmes actives, 3,5 % seront inactives l'année suivante (cf. tableau 3). La population des femmes actives étant numériquement la plus importante, ceci conduit à une baisse du taux d'activité de 0,4 point en moyenne sur la période. Plus le diplôme est élevé, plus les taux d'entrée sont élevés et les taux de sortie faibles, à l'exception notable des femmes ayant seulement un baccalauréat général. Les transitions sont plus fréquentes aux jeunes âges, que ce soit vers l'entrée ou la sortie si l'on excepte les plus de 55 ans. Ceci se traduit plutôt par un solde négatif entre 25 et 35 ans et un solde positif entre 35 et 50 ans. En termes de cycle de vie, ceci correspond à l'arrivée des jeunes enfants, qui pèse sur le taux d'activité, puis à leur croissance, qui allège les contraintes. Pour préciser cette dynamique, les transitions sont analysées selon le calendrier des naissances des enfants (4) (cf. tableau 4). Les taux de sortie sont forts à la naissance du premier enfant et encore davantage à la naissance du deuxième (5). À l'opposé, des retours à l'activité s'effectuent lorsque les enfants ont quatre ans et plus, quelle que soit la taille de la famille. L'effet désincitatif du jeune

enfant sur l'activité est donc net. En revanche, et contrairement à ce que laissent entrevoir les analyses transversales, le départ des enfants du foyer parental ne coïncide pas avec un comportement particulier de reprise d'activité.

Les variables d'activité et de fécondité évoluent donc au cours du temps dans des proportions importantes, notamment pour une même femme. Il en est de même, par construction, pour la variable d'âge, chaque femme vieillissant d'un an entre deux enquêtes consécutives. Le salaire du conjoint présente lui aussi une variabilité temporelle forte : entre deux années consécutives, seuls 60 % environ des femmes de l'échantillon voient leur conjoint demeurer dans la même tranche de salaire (cf. tableau 5). Au total, et à l'exception du diplôme, la variabilité temporelle des descripteurs est suffisamment forte pour qu'il soit possible de faire la part entre ce qui relève de leur dynamique intrinsèque et de leur caractère de substitut à l'hétérogénéité inobservable, stable au cours du temps. Pour faire la part entre ces deux interprétations possibles du rôle des variables, il est une nouvelle fois nécessaire d'avoir recours à l'économétrie, et de prendre explicitement en compte la dimension longitudinale permise par les données. Pour représenter la variété des comportements, on introduit en effet usuellement dans ce type de modèle un facteur d'hétérogénéité individuelle inobservable en plus des descripteurs habituels. Celle-ci peut d'ailleurs être mise en évidence aisément au moyen d'un premier test (cf. encadré 2). La réestimation du modèle d'activité prenant en compte une telle hétérogénéité est néanmoins plus complexe à mettre en œuvre que dans le cas où la variable expliquée est quantitative (salaire par exemple). Pour la mener à bien, on a conduit deux types d'ajustements, le premier ayant recours à un modèle *logit* à effets fixes, le second à un modèle *probit* à erreurs composées (cf. encadré 3).

Tableau 3  
**Transitions entre activité et inactivité pour les conjoints de salariés entre 1994 et 1997 selon l'âge et le diplôme**

En %

	Taux d'entrée en activité	Taux de sortie de l'activité	Accroissement
<b>Âge</b>			
Moins de 25 ans	30,0	5,5	3
De 25 à 29 ans	12,6	6,3	- 2,4
De 30 à 34 ans	10,5	4,1	- 0,6
De 35 à 39 ans	11,4	3,1	0,6
De 40 à 44 ans	9,6	1,6	1,2
De 45 à 49 ans	6,0	1,6	0,4
De 50 à 54 ans	2,5	2,3	- 0,8
55 ans et plus	0,4	10,0	- 5,1
<b>Diplôme</b>			
Pas de diplôme	3,8	4,5	- 0,4
CEP	4,4	4,5	- 1,4
CAP-BEP	7,6	3,8	- 1,1
BEPC	12,7	3,8	0,2
Baccalauréat technique	16,7	2,7	0
Baccalauréat général	8,2	3,9	- 1
Supérieur au baccalauréat	16,8	2,1	0,3
<b>Ensemble</b>	<b>7,9</b>	<b>3,5</b>	<b>- 0,4</b>

Lecture : sur 100 femmes inactives une année donnée, 7,9 deviennent actives l'année suivante ; sur 100 femmes actives une année donnée, 3,5 deviennent inactives l'année suivante ; le taux d'activité diminue alors de 0,4 point.

Source : Panel européen 1993-1997, Insee.

### À court terme, la décision d'activité dépend peu du revenu du conjoint

Les deux méthodes présentant chacune des avantages et des inconvénients différents, il a paru utile de les utiliser l'une et l'autre afin de

3. La proportion des transitions « parasites » dues à des erreurs de collecte est sans doute peu élevée dans le Panel, du fait de contrôles longitudinaux réalisés sur les variables ayant trait à l'activité.

4. On n'a cependant pas pu en dresser un bilan complet, faute d'un nombre suffisant de femmes dans certaines situations.

5. Faute d'effectifs suffisants, on ne peut analyser la naissance du troisième.

Tableau 4  
**Transitions entre activité et inactivité pour les conjoints de salariés entre 1994 et 1997  
selon la composition familiale**

En %

Année de départ	Année d'arrivée	Taux d'entrée	Taux de sortie	Accroissement
Pas d'enfant	Ensemble départ	15,5	1,6	0,5
	<i>Pas d'enfant</i>	13,6	0,9	1,1
	<i>1 enfant de moins de 4 ans</i>	25,0	5,2	- 3,2
1 enfant de moins de 4 ans	Ensemble départ	36,2	7,4	- 2,9
	<i>1 enfant de moins de 4 ans</i>	45,2	1,9	3,6
	<i>1 enfant de moins de 11 ans</i>	14,3	1,6	0,0
	<i>2 enfants dont le plus jeune a moins de 4 ans</i>	11,1	27,3	- 24,4
1 enfant de moins de 11 ans	Ensemble départ	7,3	3	- 1,6
	<i>1 enfant de moins de 11 ans</i>	9,4	1,7	- 0,4
1 enfant de plus de 11 ans dans le ménage	Ensemble départ	7,7	1,4	0,2
	<i>1 enfant de plus de 11 ans dans le ménage</i>	8,5	0,9	0,7
1 enfant hors ménage	Ensemble départ	0,0	2,9	- 2,4
	<i>1 enfant hors ménage</i>	0,0	2,4	- 2,0
2 enfants dont le plus jeune a moins de 4 ans	Ensemble départ	9,8	5,8	- 1,5
	<i>2 enfants dont le plus jeune a moins de 4 ans</i>	12,5	6,4	- 0,8
	<i>2 enfants dont le plus jeune a moins de 11 ans</i>	3,7	0,9	0,0
2 enfants dont le plus jeune a moins de 11 ans	Ensemble départ	11,9	2	0,4
	<i>2 enfants dont le plus jeune a moins de 11 ans</i>	12,6	0,7	1,4
	<i>2 enfants de plus de 10 ans dans le ménage</i>	10,5	2,8	0,0
2 enfants de plus de 10 ans dans le ménage	Ensemble départ	7,5	1,5	0,3
	<i>2 enfants de plus de 10 ans dans le ménage</i>	8,2	1,8	0,2
	<i>2 enfants de plus de 10 ans dont 1 hors ménage</i>	4,8	0,0	1,1
2 enfants de plus de 10 ans dont 1 hors ménage	Ensemble départ	0,0	3,2	- 2,7
	<i>2 enfants de plus de 10 ans dont 1 hors ménage</i>	0,0	2,3	- 1,9
	<i>2 enfants hors ménage</i>	0,0	12,9	- 10,0
2 enfants hors ménage	Ensemble départ	0,0	3,2	- 2,3
	<i>2 enfants hors ménage</i>	0,0	3,3	- 2,4
3 enfants dont le plus jeune a moins de 4 ans	Ensemble départ	9,4	10,1	2,2
	<i>3 enfants dont le plus jeune a moins de 4 ans</i>	8,3	11,9	1,6
	<i>3 enfants dont le plus jeune a moins de 11 ans</i>	15,6	6,5	4,4
3 enfants dont le plus jeune a moins de 11 ans	Ensemble départ	7,1	3,5	1,1
	<i>3 enfants dont le plus jeune a moins de 11 ans</i>	8,2	3,1	1,8
	<i>3 enfants de plus de 10 ans dans le ménage</i>	0,0	0,0	0,0
3 enfants de plus de 10 ans dans le ménage	Ensemble départ	8,2	3,0	1,0
	<i>3 enfants de plus de 10 ans dans le ménage</i>	5,7	1,8	0,8
	<i>3 enfants de plus de 10 ans dont 1 ou 2 hors ménage</i>	18,2	5,6	3,4
3 enfants de plus de 10 ans dont 1 ou 2 hors ménage	Ensemble départ	3,9	3,4	0,3
	<i>3 enfants de plus de 10 ans dont 1 ou 2 hors ménage</i>	4,0	3,6	0,4
	<i>3 enfants hors ménage</i>	4,8	3,1	0,0
3 enfants hors ménage	Ensemble départ	0,9	9,8	- 4,8
	<i>3 enfants hors ménage</i>	1,0	9,2	- 4,5

Lecture : parmi les femmes sans enfant actives une année donnée, 1,6 % sont se retirées du marché du travail l'année suivante. Si l'on se restreint aux femmes sans enfants actives mais qui ont un enfant l'année suivante, cette proportion passe à 5,2 %.

Source : Panel européen 1993-1997, Insee.

Tableau 5  
**Évolution du salaire du conjoint (en francs 1997) entre deux années consécutives**

En %

Salaire année n (1)	Salaire année n + 1 (1)								
	Chômeur	< 5 800	5 800 à 7 100	7 100 à 9 500	9 500 à 13 300	13 300 à 19 800	> 19 800	Non déclaré	Ensemble
Chômeur	63,8	14,5	7,3	4,8	4,0	0,8	0,3	4,5	100,0
< 5 800	7,8	51,2	25,4	6,7	2,6	1,7	0,9	3,7	100,0
5 800 à 7 100	3,6	12,4	54,0	23,2	3,7	1,3	0,0	1,8	100,0
7 100 à 9 500	1,2	2,5	10,9	61,8	17,2	3,7	1,1	1,6	100,0
9 500 à 13 300	1,3	1,6	1,7	13,6	65,2	13,5	2,4	0,7	100,0
13 300 à 19 800	1,0	0,9	1,4	4,9	18,3	60,7	11,8	1,0	100,0
> 19 800	1,5	0,7	0,4	2,0	7,1	16,4	70,3	1,6	100,0
Non déclaré	6,5	6,8	4,1	8,8	5,8	2,7	4,1	61,2	100,0

1. Les équivalents en euros des tranches de salaire sont donnés dans le tableau 1.

Lecture : 54 % des conjoints ayant un salaire compris entre 5 800 et 7 100 francs (884 à 1 082 €) une année ont un salaire compris entre les mêmes bornes l'année suivante.

Source : Panel européen 1993-1997, Insee.

## Encadré 2

### LES TESTS DES EFFETS INDIVIDUELS DANS LES MODÈLES LOGIT ET PROBIT

On considère un modèle de régression non linéaire de la forme :

$$y_i = f_i(\beta) + u_i \quad (1)$$

où  $\beta$  est un vecteur de paramètres et  $f_i(\beta)$  une fonction pouvant être linéaire ou non linéaire en  $\beta$ .

Cette forme fonctionnelle décrit l'espérance conditionnelle de  $y_i$  comme une fonction de  $\beta$  et de variables explicatives exogènes. Le vecteur des dérivées de  $f_i(\beta)$  par rapport aux différentes composantes de  $\beta$  est noté  $x_i(\beta)$ . Dans le cas linéaire,  $f_i(\beta)$  sera égal à  $x_i\beta$ ,  $x_i$  étant une matrice de variables explicatives ne dépendant pas de  $\beta$ .

#### La régression de Gauss-Newton

On appelle régression de Gauss-Newton une régression basée sur l'approximation du premier ordre de la série de Taylor de l'équation (1) au voisinage d'une valeur quelconque  $\hat{\beta}$  de  $\beta$  :

$$y_i - f_i(\hat{\beta}) = x_i(\hat{\beta})b + \text{résidu} \quad (2)$$

Cette régression artificielle permet d'estimer les paramètres  $b$ . Ces derniers ont une valeur non nulle, sauf si  $\hat{\beta}$  correspond aux estimateurs du modèle non linéaire (1).

Le terme de régression de Gauss-Newton provient de la méthode itérative de Gauss-Newton permettant de calculer les estimateurs des modèles non linéaires. Partant d'une valeur initiale  $\hat{\beta}$ , l'estimateur des moindres carrés ordinaires de  $b$  dans l'équation (2) fournit la direction vers laquelle  $\beta$  doit se diriger lors de l'itération suivante. On montre que la procédure de Gauss-Newton converge vers un estimateur asymptotiquement équivalent à l'estimateur des moindres carrés non linéaires  $\hat{\beta}$ . De plus, en  $\hat{\beta}$ , la matrice de variance-covariance de l'estimateur des moindres carrés ordinaires de  $b$  dans l'équation (2) peut être utilisée pour approximer la matrice de variance-covariance de  $\hat{\beta}$ .

Ces propriétés peuvent être utilisées pour réaliser des tests de spécification dans le modèle non linéaire. Supposons que l'on souhaite tester le modèle (1) contre un modèle plus général, du type :

$$y_i = f_i^*(\beta, \gamma) + u_i \quad (3)$$

où  $\gamma$  est un vecteur de paramètres, tels que, sous l'hypothèse nulle  $\gamma = 0, f_i^*(\beta, 0) = f_i(\beta)$ .

En fait, le modèle (1) apparaît comme un modèle emboîté obtenu à partir du modèle (3). Sous l'hypothèse nulle, l'estimateur du couple  $(\beta, \gamma)$  est  $(\hat{\beta}, 0)$ ; pour cette valeur des paramètres, la régression de Gauss-Newton correspondant au modèle (3) s'écrit :

$$y_i - f_i(\hat{\beta}) = x_i(\hat{\beta})b + z_i(\hat{\beta})c + \text{résidu} \quad (4)$$

où  $z_i(\hat{\beta})$  correspond à la dérivée de  $f_i^*(\beta, \gamma)$  évaluée au point  $(\hat{\beta}, 0)$ .

On peut alors montrer que les tests usuels de significativité de  $\gamma$  dans le modèle (3) sont asymptotiquement équivalents aux mêmes tests portant sur  $c$  dans le modèle (4).

#### Une application aux modèles à réponse qualitative

Cette démarche peut s'appliquer aux modèles à réponse qualitative de type *logit* ou *probit*, pour lesquels la variable observée est binaire, avec :

$$\Pr(y_i = 1) = F(x_i\beta)$$

$F$  étant la fonction de répartition d'une loi logistique ou normale.

Ce modèle peut en effet se réécrire sous la forme :

$$y_i = F(x_i\beta) + u_i \quad (5)$$

le terme d'erreur n'étant bien évidemment ni normal ni même homoscédastique.  $y_i$  correspondant à un tirage de Bernouilli dont la probabilité est  $F(x_i\beta)$ , on peut montrer que la variance de  $u_i$  vaut  $V_i = F(x_i\beta)(1 - F(x_i\beta))$ .

La régression de Gauss-Newton correspondant au modèle (5), une fois corrigée de l'hétéroscédasticité, peut alors s'écrire :

$$V_i^{-1/2}(y_i - F(x_i\beta)) = V_i^{-1/2}f(x_i\beta)x_i b + \text{résidu}$$

où  $f$  est la dérivée de la fonction de répartition  $F$ .

Cette régression permet de réaliser des tests de spécification du modèle, en particulier la nullité d'un sous-ensemble de coefficients. Comme précédemment, pour tester l'omission d'un ensemble de régresseurs  $z_i$ , il suffit de tester la significativité du coefficient  $c$  qui leur est associé dans la régression artificielle :

$$\hat{V}_i^{-1/2}(y_i - F(x_i\hat{\beta})) = \hat{V}_i^{-1/2}f(x_i\hat{\beta})(x_i b + z_i c) + \text{résidu} \quad (6)$$

Cette régression s'estime par les moindres carrés ordinaires. L'un des tests les plus adaptés est celui du score ou multiplicateur de Lagrange : il consiste simplement à calculer la somme des carrés de la partie expliquée par le modèle, et de la comparer à un  $\chi^2(p)$ , où  $p$  est le nombre de variables explicatives dans  $z$ .

Cette procédure ne présente pas véritablement d'intérêt lorsque  $p$  est petit ; il suffit en effet d'estimer les modèles (*logit* ou *probit*) contraint et non contraint, et de procéder soi-même à un test, si le logiciel utilisé ne le fournit pas en standard. En revanche, si  $p$  est très grand, comme c'est le cas en présence d'effets fixes, les logiciels courants ne permettent plus de procéder à des estimations. La méthode de la régression artificielle prend tout son intérêt. Il suffit en effet d'estimer le modèle *logit* ou *probit* sous l'hypothèse nulle, c'est-à-dire sans effets individuels (ce que réalisent tous les logiciels), et de procéder à la correction d'hétéroscédasticité, toujours sous l'hypothèse nulle. On se retrouve alors à ajuster le modèle (6), qui correspond au cas linéaire avec variables individuelles, et qui peut s'estimer en deux étapes avec le théorème de Frish-Waught. En raison de l'hétéroscédasticité, l'estimateur est néanmoins un peu plus complexe à calculer que dans le cas du modèle *within*. Le calcul du test est ensuite trivial.

### ESTIMATION DES MODÈLES LOGIT ET PROBIT À EFFETS INDIVIDUELS

Pour prendre explicitement en compte l'hétérogénéité individuelle, on dispose de deux possibilités. La première, et probablement la meilleure voie, consiste à introduire autant de variables explicatives que de femmes présentes dans l'échantillon. Ces variables sont des « effets fixes » relatifs à la personne qui perdurent au cours du temps. Leur présence dans l'estimation permet de neutraliser les corrélations apparentes entre les variables d'intérêt et les paramètres de préférences, comme dans toute estimation multivariée. L'effet fixe individuel incorpore, en outre, les variables constantes au cours du temps comme le diplôme, ce qui permet d'éviter les conséquences d'une éventuelle endogénéité du diplôme dans la décision d'activité. La difficulté est que l'approche classique en termes de maximisation de la vraisemblance ainsi construite conduit à des estimateurs non convergents lorsque la dimension temporelle est trop faible, ce qui est manifestement le cas ici. Chamberlain (1984) a proposé une procédure convergente pour  $T$  petit fondée sur une vraisemblance conditionnelle, mais elle n'est applicable que dans le cadre du modèle *logit*. Une autre voie consiste à avoir recours à un modèle à erreurs composées, dans lequel l'hétérogénéité non observable est décrite au moyen d'un terme aléatoire, dont une partie est corrélée avec les variables d'intérêt. Le test d'endogénéité des variables explicatives est alors analogue au test préconisé par Mundlak (1978) sur données quantitatives. Mais les hypothèses sont beaucoup plus restrictives que dans le cas linéaire.

#### 1. Le modèle *logit*

Le modèle *logit* avec effet individuel s'écrit :

$$P(y_{it} = 1 / x_{i1}, \dots, x_{iT}, c_i) = G(x_{it}\beta + c_i) \quad (1)$$

où est  $y_{it}$  la variable muette décrivant l'activité de l'individu  $i$  à la date  $t$ ,  $x_{i1}, \dots, x_{iT}$  les variables explicatives à chacune des dates,  $c_i$  l'effet individuel et  $G$  la fonction :

$$G(z) = \exp(z) / (1 + \exp(z))$$

L'estimation d'un tel modèle par la méthode du maximum de vraisemblance conduit à des estimateurs non convergents. Chamberlain (1984) propose une méthode fondée sur une méthode d'estimation d'une vraisemblance conditionnelle. Celle-ci s'inspire de la propriété suivante : si l'on conditionne la probabilité décrite dans (1) par le fait que la date  $t$  est la seule pour laquelle la variable d'activité vaut 1, on obtient la formule :

$$\begin{aligned} P(y_{it} = 1 / x_{i1}, \dots, x_{iT}, c_i, \sum_{t=1}^T y_{it} = 1) &= \frac{\exp(x_{it}\beta + c_i)}{\sum_{t=1}^T \exp(x_{it}\beta + c_i)} \\ &= \frac{\exp(x_{it}\beta)}{\sum_{t=1}^T \exp(x_{it}\beta)} \end{aligned}$$

L'effet individuel disparaît de la formule de la vraisemblance conditionnelle. Des calculs analogues montrent

qu'une telle propriété se généralise au conditionnement par le fait que  $s$  dates parmi  $T$  soient telles que la variable observée soit égale à 1 :

$$P(y_{it} = 1 / x_{i1}, \dots, x_{iT}, c_i, \sum_{t=1}^T y_{it} = s) = \frac{\exp(\sum_{t=1}^T x_{it}\beta y_{it})}{\sum_{d \in B_i} \exp(\sum_{t=1}^T x_{it}\beta d_t)}$$

avec :

$$B_i = \left\{ d = (d_1, \dots, d_T) / d_t \in \{0, 1\} \text{ et } \sum_{t=1}^T d_t = \sum_{t=1}^T y_{it} \right\}$$

Au total, on peut fonder l'estimation de  $\beta$  sur la maximisation de la vraisemblance conditionnelle :

$$L = \sum_{i=1}^N \log \left[ \frac{\exp(\sum_{t=1}^T x_{it}\beta y_{it})}{\sum_{d \in B_i} \exp(\sum_{t=1}^T x_{it}\beta d_t)} \right]$$

Les estimateurs obtenus sont convergents. En revanche, la procédure ne dit rien sur la valeur des effets fixes. En outre, les groupes pour les lesquels la variable expliquée est toujours nulle ou toujours égale à 1 n'interviennent pas dans l'estimation, ce qui conduit à la perte de beaucoup d'observations (on ne retient que 316 femmes, soit 11 % de l'échantillon). Cette procédure est relativement aisée à mettre en œuvre et converge très rapidement. Dans l'étude, un élément de complexité provient du fait que l'on n'a pas voulu cylindrer le modèle. Il consiste à rajouter à la log-vraisemblance des éléments analogues pour  $T = 2$  et  $T = 3$ .

#### 2. Le modèle *probit*

Le cas du modèle *logit* est particulier (il provient des propriétés des *odd-ratios*). Dans le cas général, on ne peut mettre en évidence une procédure d'estimation convergente pour  $T$  fini. Une autre façon de faire, dans le cas du modèle *probit*, est de considérer non plus un modèle à effets fixes mais un modèle à erreurs composées fondé sur la variable latente  $y_{it}^*$  :

$$y_{it}^* = x_{it}\beta + c_i + \varepsilon_{it}$$

$c_i$  est alors un effet individuel aléatoire invariant au cours du temps. Le terme  $\varepsilon_{it}$  est supposé normal (modèle *probit*), et indépendant des variables explicatives comme des effets individuels. Les aléas  $\varepsilon_{it}$  sont en outre supposés indépendants selon  $i$  et  $t$ .

L'estimation de ce type de modèle conduit cette fois à des estimateurs convergents, même pour  $T$  petit. Différentes variantes du modèle sont envisageables. La première consiste à supposer que les facteurs



s'assurer d'une certaine robustesse des conclusions. Bien que fondées sur des jeux d'hypothèses assez différents, elles aboutissent à des résultats analogues, ce qui suggère que les hypothèses effectuées étaient réalistes. En outre, les résultats sont assez différents de ceux obtenus en coupe, ce qui confirme la présence de la forte hétérogénéité déjà mise en évidence par le test. La première surprise est que le revenu du conjoint cesse d'avoir une influence significative (cf. tableau 6, modèles 2 et 3 ; tableau 7, modèle 3). Ce résultat est moins étonnant qu'il n'y paraît au vu de l'interprétation du modèle. Le nouveau modèle décompose l'effet apparent en coupe en deux effets distincts, l'un permanent absorbé par l'effet individuel et l'autre, dynamique, propre à la variable d'intérêt elle-même. D'après l'estimation, une variation du revenu courant du conjoint n'aura pas (ou aura peu) d'influence sur le comportement d'activité de la femme. En revanche, les femmes dont les conjoints ont des revenus structurellement élevés sont moins fréquemment présentes sur le marché du travail. La décision d'activité semble donc peu sensible à des variations de revenu sur le court terme, au moins quand elles sont anticipées. Pour la variable de

chômage, statut par nature peu anticipé, le coefficient est, au contraire, parmi les plus significatifs à la hausse dans les estimations longitudinales, voire quasiment significatif dans le cas du modèle *probit*. Que l'essentiel de l'influence du salaire en coupe renvoie à l'effet individuel peut s'interpréter par le fait que l'activité de la femme, engagement de long terme, se décide au vu des ressources du conjoint considérées sur une période suffisante, éventuellement le cycle de vie. Le déterminant est alors une sorte de revenu permanent, partie intégrante d'un effet individuel inobservé. Une autre interprétation consiste en la possible endogénéité de la décision de mise en couple dans l'estimation en coupe, telle qu'elle est mise en évidence par Van Der Klaauw (1996). L'échantillon des femmes vivant en couple fait alors l'objet d'une sélection endogène : parmi les femmes les plus diplômées, seules se mettent en couple celles qui sont, toutes choses égales par ailleurs, les moins enclines à l'activité. Le conditionnement par un effet individuel permettrait alors de s'affranchir de ce biais.

Un autre enseignement est la non-significativité, déjà perceptible dans le tableau 3, des

#### Encadré 3 (suite)

individuels sont des réalisations d'une variable aléatoire normale et indépendante des variables explicatives :

$$c_i = \sigma_v v_i \text{ avec } v_i \rightarrow N(0,1)$$

Le facteur aléatoire rend le calcul de la vraisemblance complexe, puisqu'il introduit un terme de couplage entre les observations se rapportant au même individu. Les aléas  $c_i$  étant par hypothèse indépendants entre eux, la vraisemblance se présente sous la forme d'un produit de termes relatifs aux seuls individus (et non aux individus et aux épisodes comme en l'absence de couplage). L'aléa individuel étant unidimensionnel, la vraisemblance peut cependant être calculée au moyen d'une procédure d'intégration numérique. L'aléa total, somme de deux aléas normaux indépendants, est lui aussi normal. Cette représentation repose cependant sur une hypothèse d'indépendance entre les facteurs individuels et les autres variables explicatives. En d'autres termes, elle ne permet pas d'évaluer, comme dans le cas du modèle à effets fixes, un biais imputable à une éventuelle corrélation entre les variables explicatives et les facteurs individuels.

Pour mettre en évidence de tels biais, Chamberlain (1984) suggère d'enrichir la spécification en supposant que l'effet individuel incorpore des informations sur les corrélations avec les variables explicatives. Plus précisément, on peut supposer, en première approximation, que la régression de  $c_i$  sur les  $x_{is}$ ,  $s = 1, \dots, T$

(notés  $x_i$ ) est linéaire avec un résidu normal et homoscédastique :

$$c_i = X_i \gamma + \sigma_w w_i, w_i \rightarrow N(0,1)$$

Le modèle à estimer s'écrit :

$$y_{it}^* = x_{it} b + x_i \gamma + \sigma_w w_i + \varepsilon_{it}$$

Techniquement, ce modèle est rigoureusement analogue au précédent si ce n'est que le nombre de variables explicatives est accru. Il s'estime au moyen des mêmes procédures.

#### 3. Limitations

Ces deux modèles, *logit* et *probit*, sont des modèles statiques. Ils supposent que, conditionnellement à l'effet individuel, la relation entre la variable expliquée et les variables explicatives est la même à toutes les dates. Par ailleurs, toujours conditionnellement à l'effet individuel, les variables explicatives sont supposées strictement exogènes, c'est-à-dire que ni leur passé ni leur futur ne dépend des chocs courants sur la variable expliquée. On ne dispose pas de procédures de tests courants permettant de valider ou d'invalider une telle hypothèse. Un signe faible de robustesse des résultats provient du fait que les estimations conduites avec les deux procédures *logit* et *probit* conduisent aux mêmes résultats, bien que ces procédures reposent sur une modélisation différente.

variables décrivant le fait que les enfants ont partiellement ou définitivement quitté le ménage. Même si les femmes dont les enfants ont quitté le foyer sont davantage actives, ce n'est pas le départ des enfants du foyer parental qui conduit la femme à se porter immédiatement sur le marché du travail. Le processus de retour à l'activité est sans doute plus diffus et plus progressif. Il est manifestement partiel, le nombre total d'enfants jouant le rôle premier sur la décision d'activité, celle-ci présentant un caractère de long terme. En revanche, l'influence négative du jeune enfant demeure. Celui-ci joue donc un rôle transitoire, rôle probablement renforcé par le dispositif réglementaire en vigueur (et notamment le succès rencontré par l'allocation

parentale d'éducation (APE)). On retrouve ici probablement les coûts liés à l'activité pour les mères de jeunes enfants mis en évidence par Lanot et Robin (1997).

Le profil d'activité par âge demeure affecté par le fait que le nombre d'enfants joue négativement en début de cycle de vie. Même en faisant apparaître un effet individuel dans l'estimation, celui-ci ne peut résoudre le problème que les enfants des plus jeunes des femmes ne sont pas encore nés. Pour pouvoir séparer chez les femmes les plus jeunes les influences respectives de la fécondité et de l'âge sur l'activité, il faudrait connaître le nombre d'enfants qu'elles auront ultérieurement. Il suffirait alors de reconstruire

Tableau 6  
Estimation des modèle *logit*, sur données empilées, en présence d'effets individuels

	Modèle 1		Modèle 2		Modèle 2bis	
	Coefficient	Student	Coefficient	Student	Coefficient	Student
Constante	2,58	22,1				
<b>Diplôme</b>						
Pas de diplôme	- 0,97	- 12,3				
CEP	- 0,47	- 5,3				
CAP-BEP	- 0,04	- 0,4				
<i>BEPC</i>	<i>Réf.</i>					
Baccalauréat technique	0,68	4,5				
Baccalauréat général	0,13	1,3				
Supérieur au baccalauréat	1,03	10,9				
Année de naissance antérieure à 1948	- 0,36	- 2,8				
<b>Âge</b>						
Moins de 25 ans	- 0,86	- 5,2	- 2,30	- 2,4	- 2,44	- 2,9
De 25 à 29 ans	- 0,30	- 2,7	- 0,46	- 0,8	- 0,62	- 1,0
De 30 à 34 ans	- 0,18	- 1,9	- 0,27	- 0,7	- 0,34	- 0,9
<i>De 35 à 39 ans</i>	<i>Réf.</i>		<i>Réf.</i>		<i>Réf.</i>	
De 40 à 44 ans	0,05	0,6	0,85	2,3	0,89	2,4
De 45 à 49 ans	- 0,16	- 1,6	1,71	2,8	1,78	3,1
De 50 à 54 ans	- 0,34	- 2,1	1,70	2,2	1,82	2,6
55 ans et plus	- 1,11	- 6,3	- 0,40	- 0,4	- 0,23	- 0,3
<b>Salaire mensuel du conjoint</b> (en francs 1997) (1)						
Conjoint chômeur	0,05	0,5	0,55	1,2		
Moins de 5 800	- 0,16	- 1,5	- 0,24	- 0,6		
De 5 800 à 7 100	- 0,11	- 1,3	- 0,24	- 0,8		
<i>De 7 100 à 9 500</i>	<i>Réf.</i>		<i>Réf.</i>			
De 9 500 à 13 300	- 0,19	- 2,5	0,34	1,2		
De 13 300 à 19 800	- 0,47	- 5,4	0,38	1,1		
Plus de 19 800	- 1,22	- 12,4	0,45	0,9		
Non déclaré	- 0,34	- 2,8	0,15	0,2		
<b>Nombre d'enfants</b>						
Aucun	0,05	0,4	- 0,60	- 0,9		
<i>Un</i>	<i>Réf.</i>		<i>Réf.</i>		<i>Réf.</i>	
Deux	- 0,64	- 8,1	- 3,99	- 6,6	- 3,88	- 6,7
Trois et plus	- 1,71	- 20,6	- 6,49	- 7,2	- 6,48	- 7,3
Enfants ayant en partie quitté le ménage	0,30	3,7	0,32	0,9		
Enfants ayant complètement quitté le ménage	1,17	9,4	0,70	1,0		
Le plus jeune enfant a moins de 4 ans	- 0,69	- 8,8	- 1,51	- 4,2	- 1,30	- 5,2

1. Les équivalents en euros des tranches de salaire sont donnés dans le tableau 1.

Lecture : les lignes marquées « Réf. » correspondent aux modalités de références des variables en italique. Les coefficients associés sont nuls afin de rendre le modèle identifiable. la première estimation (modèle 1) correspond au modèle sans effet individuel, estimé sur les données empilées. La deuxième estimation (modèle 2) correspond à la maximisation de la vraisemblance conditionnelle pour le modèle avec effets individuels. Les plages en grisé représentent les variables, constantes au cours du temps, absorbées par l'effet individuel. La troisième estimation (modèle 2bis) correspond seulement à une variante de la deuxième, avec les variables en grisé contraintes à zéro. Source : Panel européen 1993-1997, Insee.

un code de présence des enfants, qui inclurait cette fois le fait que les enfants sont à naître. Il est probable que les deux effets se superposent : les femmes qui auront davantage d'enfants auraient probablement un comportement d'activité moins intense, du fait de préférences différentes. Cependant, le fait que les enfants ne soient pas encore nés signifie que les coûts induits par leur présence future ne jouent pas encore, ce qui devrait encourager l'activité.

Tout ce raisonnement, comme toute l'étude, suppose enfin que le comportement de fécondité soit fortement exogène dans la détermination de l'activité. En d'autres termes, des chocs pré-

sents sur la variable d'activité ne doivent pas affecter le comportement de fécondité dans le passé comme dans le futur. Ce point, que l'on ne peut tester qu'au prix de la mise en œuvre de procédures très complexes, a donné lieu à de nombreux débats. Si elle est clairement erronée lorsque l'on ne conditionne pas les estimations par les effets individuels, elle est davantage admissible dans le cas contraire. On peut en effet envisager, qu'à préférences données, les décisions de fécondité soient prises en début de cycle de vie, que ce soit le nombre des enfants ou le calendrier des naissances. Auquel cas, les chocs sur la variable d'activité seraient sans impact.

Tableau 7  
Estimation des modèles *probit*, sur données empilées, en présence d'effets individuels

	Modèle 1		Modèle 2		Modèle 3	
	Coefficient	Student	Coefficient	Student	Coefficient	Student
Contante	1,51	23,1	5,44	32,4	4,89	25,6
<b>Diplôme</b>						
Pas de diplôme	- 0,58	- 12,5	- 1,45	- 16,1	- 2,31	- 21,1
CEP	- 0,28	- 5,4	- 1,05	- 10,9	- 1,05	- 9,9
CAP-BEP	- 0,02	- 0,3	- 0,98	- 11,2	- 0,32	- 3,2
BEPC	<i>Réf.</i>		<i>Réf.</i>		<i>Réf.</i>	
Baccalauréat technique	0,41	4,9	1,01	6,4	2,27	12,1
Baccalauréat général	0,07	1,1	- 0,81	- 7,6	0,19	1,6
Supérieur au baccalauréat	0,58	11,0	1,09	11,1	1,93	16,7
Année de naissance antérieure à 1948	- 0,20	- 2,6	- 2,03	- 14,3	- 1,58	- 6,0
<b>Âge</b>						
Moins de 25 ans	- 0,52	- 5,5	- 1,97	- 11,3	- 1,45	- 3,5
De 25 à 29 ans	- 0,16	- 2,6	- 0,79	- 7,1	- 0,43	- 1,5
De 30 à 34 ans	- 0,11	- 2,0	- 0,46	- 5,0	- 0,20	- 1,1
De 35 à 39 ans	<i>Réf.</i>		<i>Réf.</i>		<i>Réf.</i>	
De 40 à 44 ans	0,03	0,6	0,41	4,5	0,41	2,1
De 45 à 49 ans	- 0,09	- 1,6	0,62	5,5	0,68	2,4
De 50 à 54 ans	- 0,21	- 2,2	0,24	1,4	0,62	1,8
55 ans et plus	- 0,66	- 6,4	- 1,29	- 6,9	- 0,30	- 0,7
<b>Salaires mensuel du conjoint</b> (en francs 1997) (1)						
Conjoint chômeur	0,04	0,6	0,14	1,2	0,39	1,7
Moins de 5 800	- 0,09	- 1,5	- 0,05	- 0,5	- 0,09	- 0,5
De 5 800 à 7 100	- 0,06	- 1,2	- 0,19	- 2,1	- 0,16	- 1,0
De 7 100 à 9 500	<i>Réf.</i>		<i>Réf.</i>		<i>Réf.</i>	
De 9 500 à 13 300	- 0,10	- 2,3	- 0,08	- 1,0	0,04	0,3
De 13 300 à 19 800	- 0,27	- 5,4	- 0,19	- 2,1	0,00	0,0
Plus de 19 800	- 0,71	- 12,2	- 0,67	- 6,3	- 0,02	- 0,1
Non déclaré	- 0,21	- 2,9	- 0,25	- 1,9	0,06	0,3
<b>Nombre d'enfants</b>						
Aucun	0,02	0,3	0,07	0,5	- 0,07	- 0,2
Un	<i>Réf.</i>		<i>Réf.</i>		<i>Réf.</i>	
Deux	- 0,35	- 8,0	- 1,91	- 20,6	- 3,02	- 12,9
Trois et plus	- 0,99	- 21,2	- 3,80	- 31,4	- 4,63	- 12,7
Enfants ayant en partie quitté le ménage	0,17	3,6	0,54	6,3	0,27	1,5
Enfants ayant complètement quitté le ménage	0,67	9,2	0,91	8,3	- 0,18	- 0,7
Le plus jeune enfant a moins de 4 ans	- 0,40	- 8,7	- 1,12	- 13,5	- 0,76	- 4,8
Écart-type de l'effet individuel			3,13	40,7	3,85	36,9

1. Les équivalents en euros des tranches de salaire sont donnés dans le tableau 1.

Lecture : les lignes marquées « Réf. » correspondent aux modalités de références des variables en italique. Les coefficients associés sont nuls afin de rendre le modèle identifiable. La première estimation (modèle 1) correspond au modèle sans effet individuel, estimé sur les données empilées. La deuxième estimation (modèle 2) correspond au modèle avec effets individuels non corrélés avec les variables explicatives. La troisième estimation (modèle 3) correspond au modèle avec effets individuels corrélés avec les variables explicatives. Les coefficients correspondant aux termes linéarisés, trop nombreux, ne sont pas reproduits ici. Les coefficients des trois régressions ne sont pas directement comparables car les contraintes d'identification portant sur l'écart-type du résidu diffèrent.  
Source : Panel européen 1993-1997, Insee.

## Le coût de l'enfant pèse sur la décision d'activité

Au total, cette analyse montre une fois de plus tout l'intérêt présenté par les données longitudinales. Elles permettent d'affiner et de mieux comprendre les résultats obtenus auparavant en coupe instantanée. L'influence du revenu du conjoint est sérieusement remise en question : c'est davantage un niveau moyen sur moyenne ou longue période qui importe, en admettant que la restriction aux femmes vivant en couple n'introduise pas de biais de sélection. Les décisions en matière de fécondité importent également beaucoup, comme l'ont également montré les études antérieures. Mais, au-delà du nombre d'enfants lui-même, qui renvoie pour certaines femmes aux préférences individuelles et aux arbitrages sur le cycle de vie entre fécondité et

activité, les coûts induits par la charge des enfants comptent également pour une frange importante de femmes. Ceci transparait de deux façons : en premier lieu, le jeune enfant, plus difficile à faire garder, pèse temporairement sur l'activité des femmes. En second lieu, lorsque les enfants ont quitté le foyer parental, les taux d'activité sont plus élevés. La situation symétrique est sans doute vraie lorsque les enfants ne sont pas encore nés, mais il faudrait disposer de données longitudinales sur un horizon beaucoup plus long pour comparer les taux d'activité entre jeunes femmes selon leur descendance future. Si les choix d'activité sont obérés par la présence d'enfants à charge, les projets en matière de fécondité ne sont-ils pas alors bridés par des coûts, directs ou indirects, liés à la présence d'enfants dans le foyer parental ? □

*L'auteur remercie un lecteur d'une première version de cet article pour ses remarques et suggestions.*

---

## BIBLIOGRAPHIE

**Arellano M. et Carrasco R. (1996)**, « Binary Choice Panel Data Models with Predetermined Variables », *mimeo*, CEMFI.

**Becker G. (1985)**, « Human Capital, Effort, and the Sexual Division of Labor », *Journal of Labor Economics*, vol. 3, n° 1, pp. 33-58.

**Bourguignon F. (1986)**, « Female Participation and Taxation in France », *Unemployment, Search and Labour Supply*, Blundell and Walker eds, Cambridge University Press, pp. 243-266.

**Chamberlain G. (1984)**, « Panel Data », in Z. Griliches and M.D. Intriligator eds., *Handbook of Econometrics*, vol. II, pp. 1248-1318, Elsevier Science.

**Chiappori P.-A., Fortin B. et Lacroix G. (2001)** : « Marriage Market, Divorce Legislation and Household Labor Supply », *Journal of Political Economy*.

**Dagsvik J., Laisney F., Strom S. et Ostervold J. (1988)**, « Female Labour Supply and the Tax Benefit System in France », *Annales d'Économie et de Statistique*, n° 11, pp. 5-40.

**Eggink E., Hop J.P. et Van Praag B.M.S. (1994)**, « A Symmetric Approach to the Labour Market with the Household as a Unit of Observation », *Journal of Applied Econometrics*, vol. 9, pp. 133-161.

**Heckman J. (1974)**, « Shadow Prices, Market Wages, and Labor Supply », *Econometrica*, 42, pp. 679-694.

**Keane M. (1994)**, « A Computationally Efficient Practical Simulation Estimator for Panel Data », *Econometrica*, 62, pp. 95-116.

**Lanot J. et Robin J.-M. (1997)**, « Participation des femmes au marché du travail en présence de taxation directe et de coûts de participation », *Annales d'Économie et de Statistique*, n° 48, pp. 1-14.

**Lollivier S. (1995)**, « Activité des femmes mariées et hétérogénéité : estimation sur données de panel », *Annales d'Économie et de Statistique*, n° 39, pp. 93-106.

**Magnac T. (1997)**, « Les stages et l'insertion professionnelle des jeunes : une évaluation statistique », *Économie et Statistique*, n° 304-305, pp. 75-94.

**Mincer J. (1963)**, « Market Prices, Opportunity Costs and Income Effects », in C.F. Christ *et al.*, *Measurement in Economics*, Stanford University Press, Stanford.

**Moreau N. (2000)**, « Une application d'un modèle collectif d'offre de travail sur données françaises », *Économie et Prévision*, n° 146, 61-71.

**Mundlak Y. (1978)**, « On the Pooling of Time Series and Cross Section Data », *Econometrica*, vol. 46, n° 1, pp. 69-85.

**Riboud M. (1985)**, « An Analysis of Women's Labor Force Participation in France : Cross Section Estimates and Time-Series Evidence », *Journal of Labor Economics*, 3, n° 1, part 2, pp. 177-200.

**Van Der Klaauw W. (1996)**, « Female Labour Supply and Marital Status Decisions : a Life-Cycle Model », *Review of Economic Studies*, n° 63, pp. 199-235.

---