



Optimum Brood Size: Tests of Alternative Hypotheses

Author(s): Douglas W. Morris

Source: *Evolution*, Vol. 46, No. 6, (Dec., 1992), pp. 1848-1861

Published by: Society for the Study of Evolution

Stable URL: <http://www.jstor.org/stable/2410035>

Accessed: 06/08/2008 09:45

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=ssevol>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.

OPTIMUM BROOD SIZE: TESTS OF ALTERNATIVE HYPOTHESES

DOUGLAS W. MORRIS

*Centre for Northern Studies, Department of Biology, and School of Forestry,
Lakehead University, Thunder Bay, Ontario P7B 5E1, CANADA*

Abstract.—The most productive litter size (five) was not as common as expected in a free-living population of white-footed mice. I evaluated four competing hypotheses that can explain this pattern. Reproductive costs and annual variation in recruitment appear to be insufficient explanations for the empirical distribution of litter size. Optimal investment of reproductive resources that vary among parents is supported by some tests, but not by all. The abundance of litters less than the apparent optimum is at least partially explained by asymmetric survival in large litters (the cliff-edge hypothesis). Hypotheses that explain the empirical distribution of brood size are not mutually exclusive. Several mechanisms can act alone, or interact, to create an average brood size less than that which appears to produce the greatest number of descendants.

Key words.—Brood size, costs of reproduction, life history, litter size, *Peromyscus leucopus*.

Received April 25, 1990. Accepted April 15, 1992.

The most common clutch size in birds is often less than the size that contributes the greatest number of recruits to the population (Perrins, 1965; Klomp, 1970; De Steven, 1980; Nur, 1984; Boyce and Perrins, 1987; Møller, 1991). The same pattern has recently been reported for mammals (Morris, 1986; Boutin et al., 1988), but there are conflicting explanations for its occurrence. As outlined below, five primary hypotheses have been proposed. All assume that observed clutch or litter sizes reflect the action of natural selection maximizing individual fitness, but the hypotheses diverge in their assumptions about constraints and the character of variation in fitness in time, in space, and among individuals. In this paper I analyze juvenile recruitment in the white-footed mouse (*Peromyscus leucopus*) to assess the roles that these hypotheses play in maintaining average litter size below that which maximizes recruitment. There have been few prior attempts to determine the utility of these hypotheses in explaining litter-size patterns in natural, unmanipulated populations of mammals.

The five hypotheses, and capsule summaries of tests (emphasizing vertebrates), are as follows:

1. *The cost hypothesis* (Williams, 1966; Bell, 1984). The most productive brood size is defined to be the one that maximizes the product of the number of young produced multiplied by their own chances of survival to reproductive maturity. The optimal brood

size is lower than the most productive size when adult mortality increases with increased brood size (Charnov and Krebs, 1974). The lifetime contribution of descendants by parents producing small broods (leading to low parental mortality) may exceed that of parents sacrificing their own survival to produce large broods yielding higher recruitment.

Supporting evidence for the cost hypothesis has, at best, been equivocal (Bell and Koufopanou, 1986; Reznick, 1985; references in Boyce and Perrins, 1987; Gibbs, 1988; Pettifor et al., 1988). Though parental survivorship declines with reproductive investment in red deer (Clutton-Brock et al., 1982), this effect is not applicable to litter-size evolution because red deer produce only a single offspring per breeding season. Morris (1986) rejected the cost hypothesis by showing that maternal survival was independent of litter size in white-footed mice. The competing optimal investment hypothesis [one version of which has been labeled the individual optimization hypothesis by Pettifor et al. (1988)] challenges the importance of reproductive tradeoffs between parental survival and current fecundity in the evolutionary optimization of live histories (Ricklefs, 1968; Smith and Fretwell, 1974; Brockelman, 1975; Drent and Daan, 1980; Högstedt, 1980, 1981; Morris, 1985, 1987; Pettifor et al., 1988; Price and Liou, 1989, see below).

2. *The bad-year hypothesis* (Boyce and Perrins, 1987). If juvenile survival varies among breeding seasons and adult survivorship is low, large broods will convey high fitness when survival is high, but will have very low fitness when juvenile survival is low. Genotypes producing large broods may be drastically reduced in "bad years" and the optimal brood size may not maximize reproductive success in an average year.

No tests of the bad-year effect have yet been reported for mammals. Seasonal and annual population fluctuations of small mammals are legion and these species are likely to display temporally variable recruitment that may depend upon litter size. The bad-year effect requires more than the mere existence of variable recruitment; the variance in recruitment must be greater for large broods than it is for small ones (Boyce and Perrins, 1987).

With annual variation in recruitment, the geometric mean is a more appropriate estimate of fitness for short-lived organisms than is the arithmetic mean (Bulmer, 1985; Boyce and Perrins, 1987; Lalonde, 1991). In the great tit, Boyce and Perrins (1987) found that the clutch size that maximized geometric mean fitness was near the modal clutch size. The bad-year hypothesis thus helps to explain why modal clutch size in these passerines is less than the most productive. Assessing the generality of this explanation requires additional long-term data on recruitment in other species.

3. *The cliff-edge hypothesis* (Mountford, 1968; Boyce and Perrins, 1987). The expected number of descendants produced per brood, by a given genotype, is determined by the interaction between the frequency distribution of brood size and the proportion of offspring that are recruited. The optimal brood size will be less than the most productive whenever juvenile survival is asymmetrically low in large broods, and when the phenotypic distribution of brood sizes is either symmetrical or positively skewed. Boyce and Perrins (1987) referred to this model as the "cliff-edge effect" because the reproductive success of individuals producing large broods "falls off the cliff" relative to that of individuals producing smaller broods.

The cliff-edge hypothesis has been tested

infrequently. Consistent with the hypothesis, Morris (1986) showed that survivorship in large litters of white-footed mice was dramatically less than in smaller litters. This result was suggestive rather than conclusive; the study lasted only one year and pooled data from different habitats subsequently shown to differ in the recruitment of offspring (Morris, 1989). Boutin et al. (1988) criticized the study because the analysis implicitly assumed that an individual's survival was statistically independent of that of its littermates. Contrasting the pattern in white-footed mice, muskrat litters appear to survive or disappear as units; Boutin et al. (1988) could not find a relationship between juvenile survival and litter size in this species. The analysis below addresses this criticism with a multi-year data set corrected by the effects of habitat.

4. *The optimal investment hypothesis* (Smith and Fretwell, 1974; Brockelman, 1975; Morris, 1985, 1987). Brood size may be determined by the optimal allocation of limited resources to reproduction, and females may vary in the amount of resources available to them. Females with relatively few resources for reproduction will produce smaller broods than those with more resources. If all offspring receive similar investment, juvenile survival should be relatively invariant with brood size. Nevertheless, more recruits will tend to be produced by large than by small broods. Small broods will be more common than large ones if the average amount of resource available for reproduction to individual females is less than that necessary for the successful production of a large brood.

The optimal investment hypothesis has not been explicitly tested though the general hypothesis is supported by some data on passerine birds (Drent and Daan, 1980; Högstedt, 1980, 1981; Pettifor et al., 1988). A recent version of the model (the individual optimization hypothesis, Pettifor et al., 1988) emphasizes a reproductive tradeoff between the production of larger numbers of progeny and offspring viability. Parents in a high nutritional state, or those occupying "high quality" territories or habitats, have a larger optimum brood size than do others in the population. Individual optimization may thereby help to explain brood

size variation in species with a single brood per reproductive season.

But in many species parents face far more complicated investment decisions. The conflicting demands of lactation and gestation in mammals with postpartum estrus, for example, create a conflict for maternal resources between members of successive litters (Morris, 1986). Resolution of the conflict may result in a "solution" whereby optimum investment in a given litter, or individual, is sacrificed in the name of inclusive fitness (Morris, 1986, 1987). The term "individual optimization" is ambiguous under these circumstances (whose fitness is being maximized?), and leads to confusion whenever there is uncertainty about whether different litters, different siblings, or the parent, win the investment contest.

5. *The spatial variation hypothesis* (Perrins, 1965; Perrins and Moss, 1975; Møller, 1991; also known as the gene flow hypothesis, Dhondt et al., 1990). The optimum brood size may differ in different habitats. If individuals disperse between habitats, the optimum brood size in the population at large will be a weighted average of the optima in each habitat. Analyses confined to favorable habitats would reveal a mean brood size greater than the pooled optimum, but less than the optimum for those habitats. Analyses restricted to unfavorable habitats would reveal the opposite.

The spatial variation hypothesis has been tested on passerine bird species inhabiting woodlot islands in an agricultural landscape (Møller, 1991). The most productive clutch size was dependent upon woodlot size in five of the seven species, suggesting that optimal clutch size varied as a function of habitat. The hypothesis has also been confirmed in blue and great tits where several years' data demonstrated 1, that the optimum clutch size differed among habitats, and 2, that the observed distributions of clutch sizes were not those expected by evolution toward local optima (Dhondt et al., 1990).

Morris (1992) documented significant differences in litter size of white-footed mice occupying different habitats. Consistent with the spatial variation hypothesis, recruitment of nursing offspring into the adult population was dramatically higher in forest than in nearby fence-rows. But litters with

five offspring had higher recruitment than other litter-size classes in both habitats. The spatial variation hypothesis thus appears insufficient to account for between-habitat differences in litter size in this population.

The spatial variation hypothesis is a second-order hypothesis, for it relies on the assumption of local differences in reproductive optima without describing the primary mechanisms leading to these differences. Are they due to spatial variation in reproductive costs, in asymmetric mortality among brood sizes, in the risks of brood-size dependent mortality, in the optimal investment of resources to offspring, or to some other factor? The analyses below have been corrected for habitat-dependent differences in recruitment. This effectively eliminates the ability of the spatial variation hypothesis to account for any residual differences between mean litter size and that which is most productive.

TESTING HYPOTHESES WITH FIELD DATA

Ideally, tests of the alternative hypotheses would be based on controlled experiments that evaluate the assumptions and predictions of each hypothesis, including an elucidation of the genetic basis of phenotypic variation in brood size. Tests for reproductive costs, for example, might involve selection experiments capable of detecting antagonistic pleiotropy between current fecundity and future reproductive success (e.g., Bell and Koufopanou, 1986). In practice, there are few organisms where such detailed manipulative tests are practicable. This is particularly true for vertebrates.

Here, I am interested in assessing the utility of the hypotheses as alternative explanations for why the average litter size in white-footed mice is less than the size that produces the greatest number of recruits. I use the kind of data that can be obtained from careful demographic analyses of vertebrate populations not amenable to manipulative experiments. My tests are based on estimates of survival and fecundity of free-living individuals. The strength of my approach lies in its ability to test precise statements derived from each hypothesis on the expected relationship between brood size and subsequent reproductive success.

MATERIALS AND METHODS

Since 1981 I have monitored the life histories of white-footed mice using nest boxes located in forest, edge, and fence-row habitats in southern Ontario (Morris, 1986, 1989). White-footed mice are nocturnal, short-lived rodents. In southern Ontario, females are unlikely to produce more than six litters per year (most produce far fewer), most cease reproduction during July and August, and median longevity of adults is 171 days (Morris, 1989). Recruitment differs among habitats and between seasons (Morris, 1989). Most of my analyses, therefore, are stratified by season and correct for the influence of habitat.

The basic protocol has consisted of three nest-box checks during the spring (April–June) and two checks during the autumn (September–October) reproductive seasons. Estimates of recruitment of young from litters of known size are available from 1981 and 1983–1987 (data for 1982 are excluded because population density was extraordinarily low). Young animals were assumed to have been successfully recruited if they later occupied nest boxes as adults.

My objective is to test each hypothesis on its own merits. First, I reassess whether the most productive litter is as common in the population as expected. I then assess whether the survival of individual offspring is statistically independent of that of their siblings. I estimate the recruitment from each litter-size class to help evaluate each hypothesis.

My tests begin by comparing distributions of expected and observed numbers of recruits for different litter-size classes. The expected values are those that would be produced if the number of litters was proportional to the number of recruits born in each litter size (Fig. 1).

Following Boutin et al. (1988) I used binomial survival probabilities to evaluate whether offspring tend to survive singly or as groups. A perceptive referee noted that this procedure can yield ambiguous results because alternative protocols assess different forms of independence. An analysis of all data combined implicitly tests the hypothesis that survival rates are homogeneous among all litters. Heterogeneous sur-

vival among habitats, or among different litter sizes, could lead to rejection of the hypothesis even if the survivorship of littermates can be modeled by statistical independence. The following protocol eliminates these potentially confounding effects.

First, I calculated the total number of offspring and recruits produced by different litter sizes in each habitat and season. I then calculated the probability of any individual being recruited for each appropriate habitat and seasonal total (the ratio of total recruits to total offspring). I used these values to generate the expanded binomial probabilities of 0, 1, and ≥ 2 survivors for spring and autumn separately and for each habitat and litter size. All years were combined in this analysis because the recruitment of littermates is similar among years, see below.

The binomial survival probabilities multiplied by the number of litters of each size observed in each habitat yields the number of expected litters with different numbers of surviving recruits. These values, summed across habitats, produce the final expected distribution for different litter-size classes (litter sizes were grouped together so that the minimum expected frequency was five or greater). The expected and observed distributions were compared by a goodness-of-fit test (G -test with Williams' correction for small samples, Sokal and Rohlf, 1981, denoted by G_1 below). This analysis should reveal significant differences between the two distributions if littermates tend to survive or disappear together. A nonsignificant result implies that the survival of littermates was statistically independent and justifies the analysis of juvenile survival rates across litter-size classes. Such an analysis is essential to definitive tests of Mountford's cliff-edge hypothesis.

I also calculated the rank order of recruitment from different litter sizes across years. If survival in large litters is much less than that in smaller ones, recruitment from large litters should have a lower rank. I used these data in a supplemental test of the cliff-edge hypothesis by assessing for significant dependence of recruitment on litter-size class by Friedman's non-parametric two-way analysis of variance (Sokal and Rohlf, 1981).

I generated the expected number of recruits for different litter sizes by multiplying

habitat-specific recruitment per litter times the number of litters in each litter-size class in that habitat. I accumulated these values across habitats, pooled data to create litter-size classes with expected values greater than five, and analyzed the data by heterogeneity G -tests (denoted G_H below). This procedure allowed me to simultaneously evaluate the recruitment from different litter-size classes and the potential complication of differential responses among habitats. I repeated the heterogeneity test to search for similar effects among years. The overall analysis, if significant, would demonstrate the dependence of juvenile survival on litter size, and allow an objective assessment of the most productive litter size.

If the bad-year effect accounts for the distribution of litter sizes, the frequency of litter sizes should be proportional to the geometric mean of recruitment for each litter size. I calculated the geometric mean recruitment from all litter sizes to test this prediction.

I reassessed maternal survival costs as a function of litter size. I excluded data for 1981 that had previously rejected the cost hypothesis (Morris, 1986). Lactating females found in nest boxes at least one month beyond their date of "first capture with a litter" were classified as "survivors." I separated litters at the population median of the litter-size distribution into large (≥ 5) and small (< 5) size classes. I similarly partitioned the body size distribution of lactating females at the population median into large (≥ 93 mm) and small (< 93 mm) body sizes. I analyzed differences in survival of mothers producing large versus small litters with a hierarchical log-linear analysis (SPSS/PC+, Norusis, 1988). This procedure partitioned the likelihood-ratio chi square into additive effects to evaluate interactions between maternal body size, season of lactation, and litter size, on subsequent maternal survival. To justify my assumption of no habitat or yearly biases I repeated the analysis by substituting these effects in place of seasonal variation.

A survival cost of reproduction would be revealed in the log-linear analysis by a significant interaction between maternal survival and litter size. The potential for maternal and environmental influences to

modify this effect was analyzed by evaluating all possible interactions of third-order variables on the two-way interaction between maternal survival and litter size.

I searched for tradeoffs between current reproduction and future litter size in females that produced more than one litter. A fecundity cost (Bell, 1980) should be revealed in these females if their next observed litter tended to be smaller than their current one. I tested for this effect by correlation analysis (a few females produced litters in more than one habitat; these were excluded from the analysis).

It is possible that litter size is currently evolving to a larger size. Directional selection could compromise my ability to distinguish among the alternative hypotheses. I include a test for directional selection on litter size and a discussion of its results in the Appendix.

RESULTS

The Litter Size That Was Most Successful at Producing Recruits Was Less Common Than Expected.—Litters of size four and five were most abundant and equally common (91 and 92 litters respectively). Of the 104 known recruits, 50 were born in litters of size five, and only 25 were born in litters of size four (Fig. 1). The observed litter-size distribution thus had far fewer litters of size five than expected from the number of successful recruits ($G = 115.4$, $P \ll 0.001$, expected distribution standardized at 332 observations for the analysis of proportions).

The Survival of Littermates Appeared To Be Statistically Independent within Litter-Size Classes.—Actual recruitment per litter varied between seasons and among habitats (forest—spring = 1.0 recruits per litter, autumn = 0.29; edge—spring = 0.56, autumn = 0.07; fencerow—spring = 0.28, autumn = 0.12). The success of litters at recruiting one or more offspring (litter success), the main component of recruitment, did not vary among years (hierarchical log-linear analysis, likelihood-ratio $\chi^2 = 6.51$, $df = 5$, $P = 0.26$, 332 litters) so I combined data from all years separately for spring and autumn for each litter size in each habitat to generate the expected recruitment per litter (see Materials and Methods). The observed recruitment of littermates born in the spring

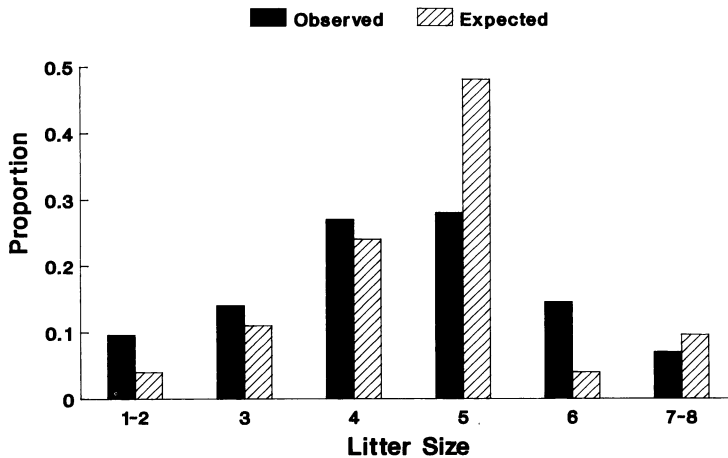


FIG. 1. The observed distribution of litter size compared to that expected if the litter-size distribution was directly proportional to the number of recruits produced by litters of different sizes. Data are presented as proportions so that the two frequency distributions can be easily compared. The graph is based on 332 litters and 104 recruits.

was consistent with the hypothesis of independent survival ($G_1 = 1.02$, hypothesis intrinsic to the data, $df = 1$, $0.1 < P < 0.5$, 161 litters, Fig. 2). A similar trend occurred in autumn but there were too few successful litters for a reliable statistical analysis (Fig. 2).

Individuals in the Largest Litters Experienced Consistently Low Survival. — Because littermates have independent survival probabilities, I calculated the expected number of recruits from spring-born litters of different sizes for each habitat. I multiplied the overall spring recruitment of the habitat by the number of offspring produced for each litter-size class, grouped these val-

ues across habitats, and compared the expected versus observed distributions with a goodness-of-fit test. Recruitment was heterogeneous across litter-size classes ($G = 12.37$, $df = 2$, $0.001 < P < 0.005$, 81 recruits, Fig. 3A). I repeated the comparison by pooling both spring and autumn data, and obtained similar results ($G = 15.62$, $df = 2$, $P < 0.005$, 104 recruits, Fig. 3B). The standardized residuals revealed that individuals in litters greater than or equal to six had consistently lower survival than individuals in smaller litter-size classes (standardized residual in spring = -2.44 , spring and autumn combined = -2.90). Litters of size five tended to have a higher rate than

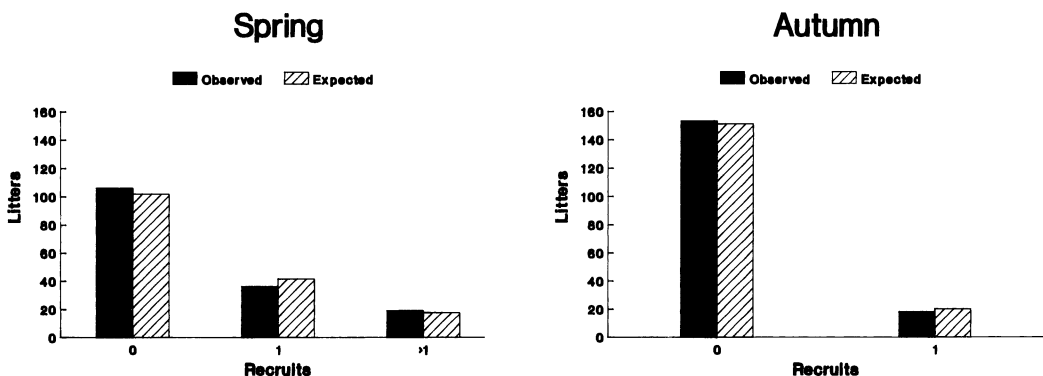


FIG. 2. Observed and expected distributions of the number of litters producing different numbers of recruits in both the spring and autumn reproductive seasons.

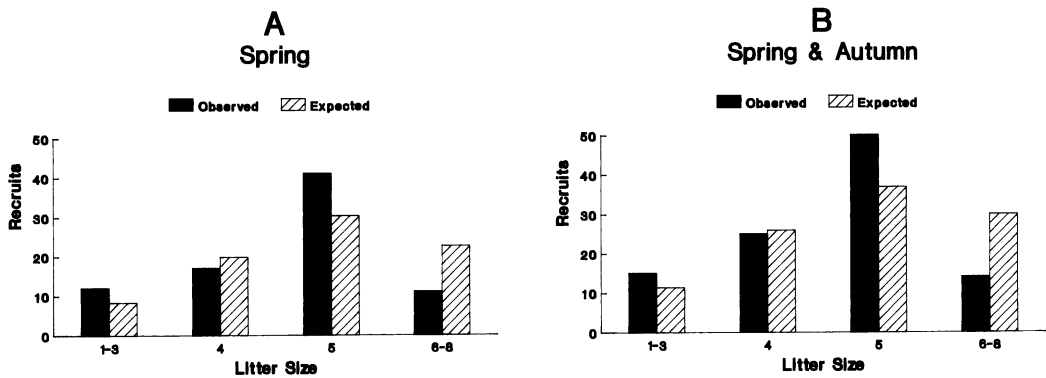


FIG. 3. The number of recruits produced by different litter-size classes. Expected values calculated relative to the total production of offspring in each litter-size class. In A, data are presented only for spring litters. In B, data are pooled for both spring and autumn reproductive seasons.

expected (standardized residual in spring = 1.94, spring and autumn = 2.17), though the difference for spring litters is only marginally significant. Survival rates for all other litter-size classes were similar (range of standardized residuals for all other litter-size classes = 1.28 to -0.63).

Litters of Size Five Contributed More Recruits Than Any Other Litter-Size Class.— Even though the recruitment from litters with five offspring was similar to that of smaller litter-size classes, litters of size five would be expected to contribute more recruits per litter than all smaller litter-size classes simply because they produced more offspring. Indeed, for spring litters, those of size five produced by far the most recruits ($G = 9.40$, $df = 2$, $0.005 < P < 0.01$, 81 recruits, Fig. 4A), independent of habitat

($G_H = 5.21$, $df = 4$, $0.1 < P < 0.5$). Pooling both spring and autumn data shows a similar effect ($G = 12.28$, $df = 2$, $0.005 < P < 0.01$, 104 recruits, Fig. 4B). Yearly variation had no residual effect on the recruitment from different litter-size classes ($G_H = 1.66$, $df = 2$, $0.1 < P < 0.5$) thus reducing the possible role of interannual variation in recruitment on litter size. An examination of the standardized residuals demonstrated, in both analyses, that litters of size five contributed significantly more recruits than expected (standardized residuals for litter-size five = 2.57 in spring, and 2.92 for the pooled data). No other litter-size class contributed significantly different numbers of recruits from those expected though all yielded fewer recruits than predicted (range of standardized residuals for all other litter-size

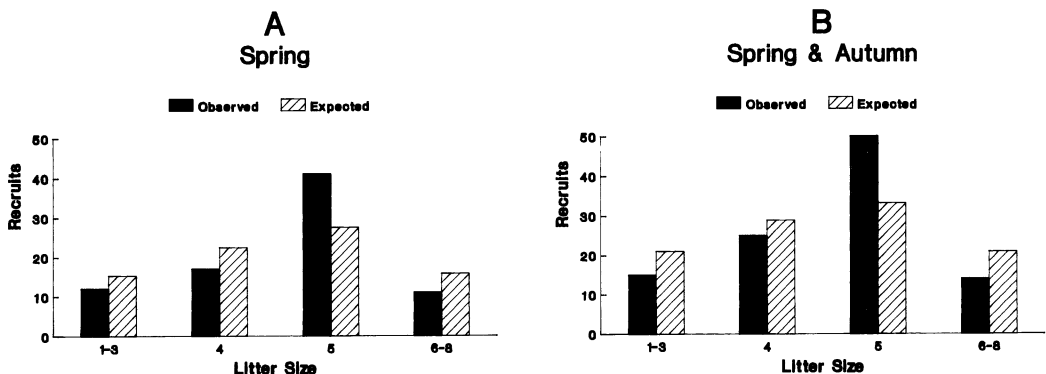


FIG. 4. Observed and expected distributions of the number of recruits produced by different litter-size classes. Expected values calculated relative to the number of recruits produced per litter. In A, data are presented only for spring litters. In B, data are pooled for both spring and autumn reproductive seasons.

TABLE 1. Recruitment per offspring of four different litter-size classes of the white-footed mouse.

Litter-size class	Year						Mean rank
	1981	1983	1984	1985	1986	1987	
1-3	0.125	0.000	0.179	0.043	0.043	0.050	2.25
4	0.115	0.067	0.100	0.056	0.031	0.012	2.50
5	0.229	0.112	0.064	0.091	0.140	0.063	1.33
6-8	0.033	0.000	0.079	0.038	0.021	0.021	3.92

classes = -1.51 to -0.73). The reduced juvenile survival in litters greater than or equal to six yielded a number of recruits per litter comparable to that produced by smaller litters.

An analysis of the ranked order of recruitment for different litter sizes across years confirmed that litters of size five had the highest recruitment, and that litters of size six or more had the lowest ($\chi^2 = 12.35$, $df = 3$, $0.005 < P < 0.01$, Table 1). These differences were not caused by differences in the ages of litters of different sizes (see Appendix).

The Geometric Mean of Recruitment Was Greatest for Litters of Size Five.—Litters of size five yielded more recruits per litter in five of the six years. But in 1984, litters of size five yielded fewer recruits than all other litter-size classes. Low recruitment in 1984 increased the variance in litters with five offspring as required by the bad-year hypothesis, but annual variation was insufficient to change the ranking of arithmetic versus geometric mean recruitment (Table 2). Geometric mean recruitment from litters of size five was more than twice as great as that of any other litter-size class. The bad-year hypothesis fails to explain why apparently sub-optimum litters of four offspring are as abundant as litters of five offspring that produce, on average, twice as many recruits.

Maternal Survival Was Independent of Litter Size.—Litter size had no discernable effect on maternal survival (survival $\times$ litter-size interaction, partial $\chi^2 = 0.22$, $df = 1$, $P = 0.64$, 261 litters, Table 3). Maternal body size influenced litter size (body size $\times$ litter-size interaction, partial $\chi^2 = 16.38$, $df = 1$, $P = 0.0001$) but had no residual effect on either maternal survival (survival $\times$ body-size interaction, partial $\chi^2 = 0.34$, $df = 1$, $P = 0.56$) or on the relationship between maternal survival and litter size (survival $\times$ litter size $\times$ body-size interaction, partial $\chi^2 = 1.64$, $df = 1$, $P = 0.20$).

Maternal survival in the spring exceeded that in the autumn (55 of 127 lactating females observed with litters in the spring were subsequently recaptured compared to 40 of the 134 similar females in the autumn, survival $\times$ season interaction, partial $\chi^2 = 5.0$, $df = 1$, $P = 0.03$), but this had no effect on survival as a function of litter size (survival $\times$ litter-size interaction, partial $\chi^2 = 0.41$, $df = 1$, $P = 0.52$, survival $\times$ season $\times$ litter-size interaction, $\chi^2 = 0.02$, $df = 1$, $P = 0.89$). Neither habitat nor annual effects had a significant influence on maternal survival (261 litters, survival $\times$ habitat interaction, partial $\chi^2 = 0.69$, $df = 2$, $P = 0.71$; survival $\times$ year interaction, partial $\chi^2 = 6.44$, $df = 4$, $P = 0.17$).

There Was No Detectable Tradeoff between Current Reproduction and Future Fe-

TABLE 2. Recruitment per litter of four different litter-size classes of the white-footed mouse.

Litter-size class	Year						Litters	AM	GM	SD
	1981	1983	1984	1985	1986	1987				
1-3	0.29	0.00	0.50	0.10	0.11	0.12	78	0.187	0.117	0.180
4	0.46	0.27	0.40	0.22	0.12	0.05	91	0.253	0.214	0.158
5	1.14	0.56	0.32	0.45	0.70	0.32	92	0.582	0.534	0.310
6-8	0.20	0.00	0.50	0.25	0.14	0.14	71	0.205	0.136	0.167
Number of litters	65	51	59	38	34	85	332			

AM—Arithmetic Mean, GM—Geometric Mean, SD—Standard Deviation.

cundity.—Forty-six of the multiple litters produced by individual females showed a change in litter size. Contrary to the cost hypothesis, successive litters tended to be larger than earlier ones: 32 were larger, 14 were smaller, and 15 were the same size. This test of potential fecundity costs may be biased. It is not obvious that a female producing a litter in the spring should be expected to incur physiological or behavioral costs lingering into the subsequent autumn.

I repeated the analysis on only those females that produced litters in consecutive nest-box checks. These lactating females are exposed to the potentially conflicting demands between nursing offspring and those developing in utero. Any detectable trade-offs between current fecundity and future within-year litter size in white-footed mice should be revealed in this analysis. The results of the analysis were opposite to expectation. Ten consecutively produced litters were of the same size, 25 were larger, and only 3 were smaller (paired *t*-test, $t = 5.09$, $df = 37$, $P < 0.0001$). The correlation between current and subsequent litter size was opposite that expected from the cost hypothesis ($r = 0.32$, $df = 36$, $P = 0.05$). These results were not caused by age differences between the two sets of litters (see Appendix).

DISCUSSION

Litters of size five were clearly more successful at contributing descendants to the population than were all other litter-size classes. Yet litters of size five were not nearly as frequent in the population as expected on the basis of their success. Which of the hypotheses account for this pattern?

The Cost Hypothesis

Previous work on this same population rejected the cost hypothesis because there was no relationship between maternal survival and litter size (Morris, 1986). The analyses presented here reinforce that conclusion and demonstrate an apparent lack of fecundity costs associated with increasing litter size.

Field tests of reproductive costs may suggest false conclusions because a covarying

TABLE 3. Maternal survival for females that produced small litters was similar to that for females that produced large litters.

Maternal survival	Litter-size class	
	<5	≥5
Recaptured	38	57
Not recaptured	74	92

trait may mask negative correlations between traits (van Noordwijk and de Jong, 1986; Bell and Koufopanou, 1986). In white-footed mice, for example, it could be argued that litter size and future survival are negatively correlated for females of a given size, but that this correlation is hidden by a positive correlation of both traits with maternal body size. The above results make this unlikely. Litter size was dependent upon maternal body size; subsequent female survival was not. It has also been suggested that environmental covariates may mask negative correlations between life history traits (Reznick, 1985). For instance, adult female mice living in one habitat could produce larger litters and live longer than females living in another. The latter argument is dispelled for these white-footed mice: adult female survival was independent of habitat.

The Assessment of Reproductive Costs

My assessment of survival cost was made by evaluating the performance of different classes of maternal phenotypes under natural conditions. The survival of a very large sample of females producing large litters was not significantly different from a similarly large sample of females producing small ones. I obtained comparable results in my assessments of fecundity costs. Subsequent litter sizes did not decline with increases in current fecundity. How can these interpretations be reconciled with the suggestion (Reznick, 1985; Bell and Koufopanou, 1986) or insistence (Lessels, 1991) that costs can only be reliably assessed by experiment? The answer lies in the intent and scale of the analysis.

The objective here is not simply to reveal costs of reproduction. The issue is to determine whether reproductive costs within a natural population of individuals are directly capable of explaining the population's frequency distribution of litter sizes. This is

a different issue than that addressed by genetic correlations that apply to the tradeoffs governing litter-size optima of genotypes.

The reason why the two approaches differ is because genotype $\times$ environment interactions can obscure reproductive costs. The existence of these interactions has frequently been used to argue that phenotypic correlations between life-history traits are unreliable indicators of genetic correlations. The door to understanding, represented by genotype $\times$ environment interactions, is locked on both sides. Genotype $\times$ environment interactions are as much of a hindrance for the geneticist, trying to explain phenotypic distributions in real populations, as they are for field biologists who wish to use phenotypic correlations as estimates of heritable covariation. The two approaches yield the same answer only when phenotypic expression is not influenced by genotype $\times$ environment interactions.

Similar issues of scale cloud our perspective on the use of field manipulations of life history in place of phenotypic correlations. Consider the case where a cost of reproduction is obscured in the phenotypic correlation by another covariate such as body size or territory quality. Whenever the covariate cannot be controlled statistically, the cost can only be revealed by experiment. Careful field manipulations of brood size could demonstrate, for example, that for females of a given body size, there is an optimum brood size. Yet even this clear demonstration of reproductive costs may be incapable of predicting the distribution of brood sizes in the population at large because that distribution is determined primarily by the distribution of female body size. In this example, reproductive costs have only an indirect influence. The correlation study would have correctly rejected the cost hypothesis as a direct functional explanation for the observed distribution of brood sizes.

My comments should not be used to argue against the existence of all reproductive costs in white-footed mice, or as an excuse to avoid experiments. I agree that life-history experiments should be conducted whenever it is reasonable and ethical to do so. But the zeal for experimentation should not be allowed to bias our judgments on the

utility of other kinds of evidence. This latter point is especially crucial when alternative approaches address different issues.

The Bad-Year Hypothesis

The bad-year effect, by itself, is incapable of accounting for why the mean litter size is substantially less than the size with the greatest recruitment. The geometric mean recruitment of one of the two modal litter sizes (four) was substantially below the other that was most productive (five).

Nevertheless, this population of white-footed mice occasionally experiences tragically bad years. Mortality was high during the late autumn of 1981 and the subsequent winter. Most boxes were empty by spring 1982 and I observed only four litters during that season, one each of size four and six, and two of size five. Two of these litters contributed a single recruit, one from litter-size four and the other from litter-size five. It is thus possible that I have underestimated the potential of annual variation in recruitment to account for the empirical distribution of litter sizes. There are too few data to tell even after six years of study.

The Cliff-Edge Hypothesis

Data published from the spring of 1981 supported the cliff-edge hypothesis. The recruitment from litters of size six and greater was much less than that for smaller litter-size classes (Morris, 1986). Now, with six years of data, the evidence in favor of the cliff-edge hypothesis is even more compelling. The chances of recruitment per offspring have remained consistently low for large litters (six to eight offspring). This view must be balanced, however, by the observation that large and small litters (fewer than five offspring) produce, on average, similar numbers of recruits. It may be prudent to postpone final judgment on the cliff-edge hypothesis until we have a reasonable explanation for the proximate mechanism producing low recruitment in large litters. The optimal investment hypothesis provides one such mechanism (Morris, unpubl. data).

The Optimal Investment Hypothesis

If all parents allocate resources similarly, parents with relatively few resources avail-

able for reproduction should produce smaller broods than those with more resources. Brood size in magpies, for example, is tightly linked to territory quality (Högstedt, 1980, 1981) and the distribution of brood sizes and subsequent recruitment reflects the availability of territories of differential quality.

Female white-footed mice are also territorial during the breeding season (Burt, 1940; Nicholson, 1941; Stickel, 1968; Metzgar, 1971; Rowley and Christian, 1976). Variation in territory quality could account for differences in litter size and recruitment. If so, some breeding sites (boxes) should consistently "produce" more recruits than others. My observations falsify this prediction. The accumulated distribution (controlled by habitat) of several years' data on litter success in different nest boxes is not significantly different from random expectation (Morris, 1991).

The data on recruitment are, at first glance, also inconsistent with the optimal investment argument. If all offspring receive a similar investment, regardless of litter size, and if juveniles survive independently of one another, the chances of an offspring's recruitment should be similar in litters of different sizes. My data on white-footed mice show a clear advantage for litters of size five. But the assumption of similar investment for all offspring of a given litter size may be unreasonable when parents err in their investment. In white-footed mice, for example, large females produce larger litters than do small females (Morris, 1992, unpubl. data). Yet some small females produce large litters, and some large females produce small ones. Offspring in large litters produced by small females have a significantly lower chance of recruitment than offspring in large litters produced by large females (Morris, unpubl. data). These results are consistent with optimal investment. The hypothesis also predicts that litter size should be greater in habitats with the greatest survival risks (Morris, 1987), a result confirmed in this population of white-footed mice (Morris, 1992).

*The Hypotheses Are Not Mutually
Exclusive*

Whatever the various proximate mechanisms it is apparent that, over the six years

of this study, phenotypic selection has favored litters of size five over all other litter-size classes. It is also apparent that the empirical distribution of litter sizes is markedly different from that imposed by selection on these intermediate-sized litters. Of the hypotheses capable of accounting for this pattern, only the cost and bad-years hypotheses have been rejected. Phenotypic correlations that fail to detect a cost of reproduction are consistent with the alternative of optimal investment. The observed cliff-edge effect may also be a special case of optimal investment (Morris, unpubl. data). The lesson for students of life history is, quite appropriately, one of compromise. The hypotheses are not mutually exclusive, and indeed may sometimes be hierarchically dependent. All may contribute independently or synergistically, at different times, and at different scales, to the persistence of a smaller mean brood size than the size that appears to contribute the most descendants to future generations.

ACKNOWLEDGMENTS

I thank K. Morris whose field and computer assistance were essential to this project. S. Boutin graciously sent me a stimulating manuscript that demonstrated the need to test for nonindependent survival of littermates. I thank R. Holt for a thorough review and numerous editorial suggestions that resulted in marked improvements in both style and substance. Candid comments by J. Hutchings and anonymous referees helped me to improve the paper. I also thank J. Barten, C. Learn, and E. Gajda for collecting field data in my absence, J. Enright for frequent field assistance, P. Anderson for providing nest-box lumber and early encouragement, Monarch Mattress Company for free and copious nesting material, and Morridell Farms Ltd. and R. Wylie for access to research sites. The 1981 field work was completed while I held a NSERC post-doctoral fellowship at The University of Western Ontario. The continued support of Canada's Natural Sciences and Engineering Research Council (grant #OGP0116430) is gratefully appreciated.

LITERATURE CITED

- BELL, G. 1980. The costs of reproduction and their consequences. *Am. Nat.* 116:45-76.
———. 1984. Measuring the cost of reproduction. I.

- The correlation structure of the life table of a plankton rotifer. *Evolution* 38:300-313.
- BELL, G., AND V. KOUFOPOULOU. 1986. The cost of reproduction, pp. 83-131. In R. Dawkins (ed.), *Oxford Surveys of Evolutionary Biology*, Vol. 3. Oxford University Press, Oxford, UK.
- BOUTIN, S., R. A. MOSES, AND M. J. CALEY. 1988. The relationship between juvenile survival and litter size in wild muskrats (*Ondatra zibethicus*). *J. Anim. Ecol.* 57:455-462.
- BOYCE, M. S., AND C. M. PERRINS. 1987. Optimizing great tit clutch size in a fluctuating environment. *Ecology* 68:142-153.
- BROCKELMAN, W. Y. 1975. Competition, fitness of offspring, and optimal clutch size. *Am. Nat.* 109: 677-699.
- BULMER, M. G. 1985. Selection for iteroparity in a variable environment. *Am. Nat.* 126:63-71.
- BURT, W. H. 1940. Territorial behavior and populations of some small mammals in southern Michigan. *Misc. Publ. Mus. Zool. Univ. Mich.* 45:1-58.
- CHARNOV, E. L., AND J. R. KREBS. 1974. On clutch size and fitness. *Ibis* 116:217-219.
- CLUTTON-BROCK, T. H., F. E. GUINNESS, AND S. D. ALBON. 1982. *Red Deer: The Behavior and Ecology of Two Sexes*. University of Chicago Press, Chicago, IL USA.
- COOKE, F., P. D. TAYLOR, C. M. FRANCIS, AND R. F. ROCKWELL. 1990. Directional selection and clutch size in birds. *Am. Nat.* 136:261-267.
- DE STEVEN, D. 1980. Clutch size, breeding success, and parental survival in the tree swallow (*Iridoprocne bicolor*). *Evolution* 34:278-291.
- DHONDT, A. A., F. ADRIANSEN, E. MATTHYSEN, AND B. KEMPENAEERS. 1990. Nonadaptive clutch sizes in tits. *Nature* 348:723-725.
- DRENT, R. H., AND S. DAAN. 1980. The prudent parent: Energetic adjustment in avian breeding. *Ardea* 68:225-252.
- GIBBS, H. L. 1988. Heritability and selection on clutch size in Darwin's medium ground finches (*Geospiza fortis*). *Evolution* 42:750-762.
- HÖGSTEDT, G. 1980. Evolution of clutch-size in birds: Adaptive variation in relation to territory quality. *Science* 210:1148-1150.
- . 1981. Should there be a positive or negative correlation between survival of adults in a bird population and their clutch size? *Am. Nat.* 118:568-571.
- KLOMP, H. 1970. The determination of clutch size in birds: A review. *Ardea* 58:1-124.
- LALONDE, R. G. 1991. Optimal offspring provisioning when resources are not predictable. *Am. Nat.* 138:680-686.
- LESSELS, C. M. 1991. The evolution of life histories, pp. 32-65. In J. R. Krebs, and N. B. Davies (eds.), *Behavioural Ecology: An Evolutionary Approach*, 3rd ed. Blackwell Scientific Publications, London, UK.
- METZGAR, L. H. 1971. Behavioral population regulation in the woodmouse. *Peromyscus leucopus*. *Am. Midl. Nat.* 86:434-447.
- MØLLER, A. P. 1991. Clutch size, nest predation, and distribution of avian unequal competitors in a patchy environment. *Ecology* 72:1336-1349.
- MORRIS, D. W. 1985. Natural selection for reproductive optima. *Oikos* 45:290-292.
- . 1986. Proximate and ultimate controls on life-history variation: The evolution of litter size in white-footed mice *Peromyscus leucopus*. *Evolution* 40:169-181.
- . 1987. Optimal allocation of parental investment. *Oikos* 49:332-339.
- . 1989. Density-dependent habitat selection: Testing the theory with fitness data. *Evol. Ecol.* 3:80-94.
- . 1991. Fitness and patch selection by white-footed mice. *Am. Nat.* 138:702-716.
- . 1992. Environmental networks, compensating life histories, and habitat selection by white-footed mice. *Evol. Ecol.* 6:1-14.
- MOUNTFORD, M. D. 1968. The significance of litter-size. *J. Anim. Ecol.* 37:363-367.
- NICHOLSON, A. J. 1941. The homes and social habits of the wood-mouse (*Peromyscus leucopus noveboracensis*) in southern Michigan. *Am. Midl. Nat.* 25: 196-223.
- NORUSIS, M. J. 1988. *SPSS/PC+ Advanced Statistics V2.0*. SPSS Inc., Chicago, IL, USA.
- NUR, N. 1984. The consequences of brood size for breeding blue tits. I. Adult survival, weight change, and the cost of reproduction. *J. Anim. Ecol.* 53: 479-496.
- PERRINS, C. M. 1965. Population fluctuations and clutch size in the Great Tit *Parus major* L. *J. Anim. Ecol.* 34:601-647.
- PERRINS, C. M., AND D. MOSS. 1975. Reproductive rates in the great tit. *J. Anim. Ecol.* 44:695-706.
- PETTIFOR, R. A., C. M. PERRINS, AND R. H. MCCLEERY. 1988. Individual optimization of clutch size in great tits. *Nature* 336:160-162.
- PRICE, T., M. KIRKPATRICK, AND S. J. ARNOLD. 1988. Directional selection and the evolution of breeding date in birds. *Science* 240:798-799.
- PRICE, T., AND L. LIQU. 1989. Selection on clutch size in birds. *Am. Nat.* 134:950-959.
- REZNICK, D. 1985. Costs of reproduction: An evaluation of the empirical evidence. *Oikos* 44:257-267.
- RICKLEFS, R. E. 1968. On the limitation of brood size in passerine birds by the ability of adults to nourish their young. *Proc. Nat. Acad. Sci. USA* 61:847-851.
- ROCKWELL, R. F., C. S. FINDLAY, AND F. COOKE. 1987. Is there an optimal clutch size in snow geese? *Am. Nat.* 130:839-863.
- ROWLEY, M. H., AND J. J. CHRISTIAN. 1976. Intraspecific aggression of *Peromyscus leucopus*. *Behav. Biol.* 17:249-253.
- SMITH, C. C., AND S. D. FRETWELL. 1974. The optimal balance between size and number of offspring. *Am. Nat.* 108:499-506.
- SOKAL, R. R., AND F. J. ROHLF. 1981. *Biometry*, 2nd ed. Freeman, San Francisco, CA USA.
- STICKEL, L. F. 1968. Home range and travels, pp. 373-411. In J. A. King (ed.), *Biology of Peromyscus (Rodentia)*. American Society of Mammalogists, Stillwater, OK USA.
- VAN NOORDWIJK, A. J., AND G. DE JONG. 1986. Acquisition and allocation of resources: Their influence on variation in life history tactics. *Am. Nat.* 128:137-142.
- VAN NOORDWIJK, A. J., J. H. VAN BALEN, AND W. SCHARLOO. 1981. Genetic and environmental

variation in clutch size of the great tit (*Parus major*). *Neth. J. Zool.* 31:342–372.

WILLIAMS, G. C. 1966. Natural selection, the costs of reproduction, and a refinement of Lack's hypothesis. *Am. Nat.* 100:687–690.

Corresponding Editor: R. Holt

APPENDIX

Natural Selection on Litter Size

In an attempt to better understand the potential effect of natural selection on litter size I estimated annual selection differentials. I calculated the difference between the average litter size and the average size of those litters that were successful at recruiting offspring, weighted by the number of recruits (van Noordwijk et al., 1981; Boyce and Perrins, 1987).

The Data Were Consistent with Directional Selection Whereas Evidence for Stabilizing Selection Was Equivocal.—The mean litter size tended to be smaller than the weighted mean size of successful litters, but differences within years were not statistically significant (Table A1). I evaluated the trend for mean litter sizes to be larger in successful litters than in unsuccessful litters by a paired *t*-test across years. The analysis was significant ($t = 2.49$, $df = 5$, $P = 0.03$, one-tailed test) suggesting that directional selection may be acting to increase litter size, but see below.

I assessed the potential for stabilizing selection by comparing the coefficients of variation of litter size of all litters observed with those that were successful at recruiting offspring. This estimate tended to be smaller in successful litters, but the difference was not statistically significant (paired *t*-test, $t = 2.43$, $df = 5$, $P = 0.06$, two-tailed test). The results were similar in a comparison of coefficients of variation between successful and unsuccessful litters ($t = 2.27$, $df = 5$, $P = 0.07$). Analyses on only spring litters gave comparable results. In all cases, however, the coefficient of variation in litter size was larger following selection for only one of the six yearly comparisons (1984).

Directional selection toward a larger litter size may explain the distribution of litter sizes. The data clearly demonstrate that, during spring, litters of size five produce more recruits than any other litter size. The disproportionate number of offspring contributed to the population by these litters results in positive selection differentials, and phenotypic selection for increased litter size.

A positive correlation between clutch size and recruitment in recent avian studies (Boyce and Perrins, 1987; Rockwell et al., 1987; Gibbs, 1988) is also consistent with directional selection toward increased brood size. But positive selection differentials need not imply selection, even if there is additive genetic variance for clutch size. In one view, nonheritable variation (such as the health or nutritional state of the parent) may affect both the number of young produced as well as the total number recruited (Price et al., 1988). Parents in a high nutritional state may produce large broods with high juvenile survival. Such a process can create positive selection differentials, but if nutritional state is nonheritable, mean brood size will remain unchanged between generations (Price and Liou, 1989). Alternatively, selection may be operating on the ability of parents to compete for territories or resources (Cooke et al., 1990). Parents in high quality territories (or parents in a high nutritional state) may produce larger broods and more descendants compared to parents in low quality territories. If the number of territories (resource abundance) is limited (and stable), there will be no selection toward increased brood size even though average competitive ability has improved (Cooke et al., 1990).

Early Juvenile Mortality, Litter Size and Recruitment

My field protocol resulted in litters being discovered at different ages. Older litters may have suffered some mortality prior to my observations. Death of littermates in large litters could create an artifact of high recruitment in intermediate-sized litters. Many litters that I recorded with five young, for example, may have originally been composed of six or more offspring, and any successful recruitment from such litters would inflate my estimate of recruitment from litters of size five. I tested for this effect by comparing the mean age of litters with five young against that for litters with six or more offspring. In anticipation of a significant age effect on litter size I eliminated all litters except those first observed with immature young (<14 days old) and repeated the recruitment analysis. These tests were followed by a similar evaluation of the potential for age-related differences in litter size to obscure fecundity costs.

Differences in Chances of Recruitment Per Offspring in Large Litters Were Not Caused by Differences in the Ages of Litters of Different Sizes.—The estimated age of litters with 5 offspring was significantly greater than that of larger litters (mean of litters of size 5 = 12.83

TABLE A1. Selection differentials for white-footed mouse litter size in six different years.

Year	Litters	Recruits	Observed litters		Successful litters		Selection differential
			Mean	CV	Mean	CV	
1981	65	34	4.11	32.52	4.44	24.90	0.33 NS
1983	51	13	4.55	28.93	4.69	10.23	0.14 NS
1984	59	24	4.64	25.54	4.54	31.79	−0.10 NS
1985	38	10	4.37	35.65	5.00	24.94	0.63 NS
1986	34	10	4.47	34.94	4.90	20.29	0.43 NS
1987	85	13	4.41	34.68	4.92	29.27	0.51 NS

CV—Coefficient of Variation, NS—No significant difference in mean litter size.

TABLE A2. Recruitment for litter-size classes five and six or greater (aged 13 days or less).

Litter-size class	Number of litters	Recruits per litter	Number of offspring	Recruits per offspring
5	41	0.39	270	0.06
6-8	35	0.09	331	0.01

days, mean of litters of size 6-8 = 10.36 days, $t = 2.14$, $df = 188$, $P = 0.034$). My recruitment estimates for litters of size five could, therefore, be inflated because some of the successful litters that I observed with five offspring may have represented the survivors of litters originally composed of six or more offspring. This could result in some large litters being misclassified into smaller litter-size classes simply because they were discovered after a period of early juvenile attrition.

I repeated the recruitment analysis on only immature litters (13 days of age or less). The advantage of litters of size five was equally dramatic (Table A2) even though there was no significant difference in the mean age of litters in the two size classes [mean = 6.2 days for litters of size five, mean = 5.2 days for litters of six or more, $t = 1.39$, $df = 104$, $P = 0.17$, analysis includes litters that were too young to eartag, results were similar for the comparison of marked immature litters only ($t =$

0.77, $df = 74$, $P = 0.44$)). Any misclassification of large litters is unlikely to be responsible for the high recruitment from litters of size five.

The Apparent Absence of Fecundity Costs was not Caused by Age Differences between Litters.—I was concerned that the overwhelming evidence for an increase in litter sizes observed in my test of fecundity costs might be related to age differences between consecutively produced litters. If consecutive litters tend to be observed at a younger age than the previous litter, and if my estimates of litter size include attrition due to age-specific mortality since birth, my data on fecundity costs could be biased in favor of a negative result. As anticipated, consecutive litters were younger than the previous litter (25 of 38 cases, mean ages of 12.3 and 8.3 days respectively, paired t -test, $t = 3.6$, $df = 37$, $P = 0.0009$).

Was the significant difference in the age of consecutively produced litters sufficient to account for the dramatic increase in litter size? I analyzed for differences in the size of previous and consecutive litters while controlling for their age as a covariate in an analysis of covariance design. Neither of the within-group regressions relating litter size to age were statistically significant ($F = 1.91$, $df = 1,36$, $P = 0.18$, and $F = 0.74$, $df = 1,36$, $P = 0.40$). The increase in litter size in consecutively produced litters was not caused by differences in their respective ages.